

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2006, 17:51:20 ; Search time 10461 Seconds  
(without alignments)  
17149.214 Million cell updates/sec

Title: US-10-681-086-1

Perfect score: 3156

Sequence: 1 ggatccacagggtacgac.....atcggtatgatgcgaattc 3156

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb\_ba.\*

2: gb\_in.\*

3: gb\_env.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pr.\*

9: gb\_ro.\*

10: gb\_sts.\*

11: gb\_sy.\*

12: gb\_un.\*

13: gb\_vi.\*

14: gb\_htg.\*

15: gb\_pl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result/No.	Score	Query Match	Length	ID	Description
1	3156	100.0	3156	6	BD175786
2	3156	100.0	3156	6	AR435692
3	3156	100.0	3156	6	AX394258
4	1612.6	51.1	8478	6	AX173058
5	1612.6	51.1	8478	6	AR630545
6	1612.6	51.1	10153	1	BSU51868
7	1612.6	51.1	199922	1	BSUB0016
8	1612.6	51.1	220060	1	AF008220
9	1431	45.3	6739	1	AB088066
10	1219	38.0	3560	6	AX825002
11	1200.4	38.0	7350	6	CQ894464
12	1192.8	37.8	8152	6	CS115035
13	1192.8	37.8	8644	6	CS115022
14	1192.8	37.8	13129	6	CS091382
15	867	27.5	4335	6	CS119875
16	867	27.5	4545	1	PB0KANRCG
17	867	27.5	4548	1	PB0110CG
18	867	27.5	4780	11	ACVPBD64

19	867	27.5	5115	1	AB037420	AB037420 Staphyloc
20	867	27.5	5142	6	AX069289	AX069289 Sequence
21	867	27.5	5793	6	AR569128	AR569128 Sequence
22	867	27.5	5943	6	AR569127	AR569127 Sequence
23	867	27.5	6661	6	AX601433	AX601433 Sequence
24	867	27.5	7311	6	CS091381	CS091381 Sequence
25	867	27.5	7336	6	A08030	A08030 Synthetic n
26	867	27.5	7336	6	A13198	A13198 complete pl
27	867	27.5	7336	6	A20502	A20502 Hind III fr
28	867	27.5	7336	6	AR033984	AR033984 Sequence
29	867	27.5	7336	6	AR123993	AR123993 Sequence
30	867	27.5	7336	6	AR154323	AR154323 Sequence
31	867	27.5	7336	6	AR408644	AR408644 Sequence
32	867	27.5	8119	6	AR002333	AR002333 Sequence
33	867	27.5	8119	6	AR018032	AR018032 Sequence
34	867	27.5	8119	6	AR055698	AR055698 Sequence
35	867	27.5	9144	6	CQ815355	CQ815355 Sequence
36	867	27.5	46445	1	AF051917	AF051917 Staphyloc
37	867	27.5	58237	1	D86934	D86934 Staphylococ
38	867	27.5	110000	1	BA000017	BA000017 Staphyloc
39	867	27.5	110000	1	BA000018	BA000018 Staphyloc
40	867	27.5	110000	1	BX571856	BX571856 Staphyloc
41	865.4	27.4	1200	1	PTBKAN	K02551 plasmid pTB
42	865.4	27.4	2262	1	BSPRBHJK	X03409 Thermophil
43	865.4	27.4	4525	1	BSREPB	X15670 Bacillus pl
44	865.4	27.4	5767	6	AX035965	AX035965 Sequence
45	865.4	27.4	6447	11	AF288421	AF288421 Synthetic

#### ALIGNMENTS

RESULT 1  
BD175786  
LOCUS BD175786 3156 bp DNA linear PAT 18-MAR-2003  
DEFINITION Process for producing target fermentation product.  
ACCESSION BD175786  
VERSION BD175786.1 GI:29121488  
KEYWORDS JP 2002253273-A/1.  
SOURCE Bacillus subtilis  
ORGANISM Bacillus subtilis  
REFERENCE 1 (bases 1 to 3156)  
AUTHORS Hohmann,H.P., Mouncey,N.J., Schlieker,H.W. and Stebbins,J.W.  
TITLE Process for producing target fermentation product  
JOURNAL Patent: JP 2002253273-A 1 10-SEP-2002;  
ROCHE VITAMINS AG  
COMMENT OS Bacillus subtilis  
PN JP 2002253273-A/1  
PD 10-SEP-2002  
PF 08-AUG-2001 JP 2001240984  
PR 08-AUG-2000 US 09/633927  
PI HANS PETER HOHMANN,NIGEL JOHN MOUNCEY,HEINRICH WINFRIED PI  
SCHLIEKER,  
PI JEFFREY W STEBBINS  
PC C12N15/09,C12N1/21,C12P25/00//C12N1/21,C12R1:125), (C12P25/00,  
PC C12R1:125),  
PC C12N15/00,  
CC Process for producing target fermentation product PH Key  
Location/Qualifiers  
FT source  
FT 1. 3156  
Location/Qualifiers  
1. 3156  
/organism='Bacillus subtilis'.  
Location/Qualifiers  
1. 3156  
/organism='Bacillus subtilis'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:1423'

Query Match 100.0%; Score 3156; DB 6; Length 3156;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GGATCCACGAGGTTACGAGCCTTGAAGATTGATTTCTCGGTTAAACGAGCGGTTAGACAGA	60
Db	1	GGATCCACGAGGTTACGAGCCTTGAAGATTGATTTCTCGGTTAAACGAGCGGTTAGACAGA	60
Qy	61	ATGAAAGAACGGCGGTACATCGTAACCTGCGGTCAATGATGAGCGCGGTTCCAGAG	120
Db	61	ATGAAAGAACGGCGGTACATCGTAACCTGCGGTCAATGAGATGAGCGCGGTTCCAGAG	120
Qy	121	AGGAATATTGATGGCGAAAAATCAACCGGTCCTGGTCCTCAAAACAATTTATTAGGCGTCGCA	180
Db	121	AGGAATATTGATGGCGAAAAATCAACCGGTCCTGGTCCTCAACAATTTATTAGGCGTCGCA	180
Qy	181	AGCGATAGACGTTTGATCGATGCGAGCCCAACAGCAATTCGACGCAATTTGGGACAGGAAGC	240
Db	181	AGCGATAGACGTTTGATCGATGCGAGCCCAACAGCAATTCGACGCAATTTGGGACAGGAAGC	240
Qy	241	AGCGGTTACCGTTTAAACGACAGGCAATTCGGTCTGGCATGAAAGCTAGAAAGAAGATT	300
Db	241	AGCGGTTACCGTTTAAACGACAGGCAATTCGGTCTGGCATGAAAGCTAGAAAGAAGATT	300
Qy	301	GCCAGCTTTAAACTGACAGAAAGCGGCCCTGCTGTTTTCGAGCGGTACTTCGCCCAATGTC	360
Db	301	GCCAGCTTTAAACTGACAGAAAGCGGCCCTGCTGTTTTCGAGCGGTACTTCGCCCAATGTC	360
Qy	361	GGTGTCTTTTCATCTCTGCCAGAAAAGGAAGATGTCAATTTAAAGTGCACGCTCAATCAT	420
Db	361	GGTGTCTTTTCATCTCTGCCAGAAAAGGAAGATGTCAATTTAAAGTGCACGCTCAATCAT	420
Qy	421	GCAAGTATGATCGAGCGGTGCGGACTTTCCTAAGCGTGATACAGTTGTTTATCGGCATATT	480
Db	421	GCAAGTATGATCGAGCGGTGCGGACTTTCCTAAGCGTGATACAGTTGTTTATCGGCATATT	480
Qy	481	GATATGAATGATCTTTGAAAAACAAGCTGAATGAAAAACAGCGGTATACAGCGCGGTTTATC	540
Db	481	GATATGAATGATCTTTGAAAAACAAGCTGAATGAAAAACAGCGGTATACAGCGCGGTTTATC	540
Qy	541	GTAACAGACGGAGTATTACAGATGGATGGCAAAATCGCCCCCTCTTGATCAGATCATCTCA	600
Db	541	GTAACAGACGGAGTATTACAGATGGATGGCAAAATCGCCCCCTCTTGATCAGATCATCTCA	600
Qy	601	CTTCGGAAACGCTATATGCGCTTCGTGGTTCGTTGATGATGCCACAGGAGTTTG	660
Db	601	CTTCGGAAACGCTATATGCGCTTCGTGGTTCGTTGATGATGCCACAGGAGTTTG	660
Qy	661	GGCGATTTCGGGACAAAGGAAACAGTGAATACATTTCGTGTTGTCGCCGACATTGTTATCGGC	720
Db	661	GGCGATTTCGGGACAAAGGAAACAGTGAATACATTTCGTGTTGTCGCCGACATTGTTATCGGC	720
Qy	721	ACCTTAAGCAAAAGCTGTTGGCGCGGAAAGGAGTTTTCGGCAGGATCAGCGGCTCTTCATC	780
Db	721	ACCTTAAGCAAAAGCTGTTGGCGCGGAAAGGAGTTTTCGGCAGGATCAGCGGCTCTTCATC	780
Qy	781	GACTTTTTGTGAACCATGCGCAGAACATTTATCTTTCAAAACCGGTATTCGCCAGCCAGC	840
Db	781	GACTTTTTGTGAACCATGCGCAGAACATTTATCTTTCAAAACCGGTATTCGCCAGCCAGC	840
Qy	841	TGTCGGCTGCTCACGAGGCTTTTCAAACATCATTTGAAGCCAGCGGAAAAACGACAGCTT	900
Db	841	TGTCGGCTGCTCACGAGGCTTTTCAAACATCATTTGAAGCCAGCGGAAAAACGACAGCTT	900
Qy	901	TTATTTTCTTATATCAGCATGATCAGAAACCAAGTCTGAAGAAATATGGGTTATGTGGTGA	960
Db	901	TTATTTTCTTATATCAGCATGATCAGAAACCAAGTCTGAAGAAATATGGGTTATGTGGTGA	960
Qy	961	GGAGATCACACCGATTTATTCCTGTAGTCAATTGGCGGATGCCCATAAAAACGCTCCTATT	1020
Db	961	GGAGATCACACCGATTTATTCCTGTAGTCAATTGGCGGATGCCCATAAAAACGCTCCTATT	1020
Qy	1021	GCTGAAAAACCTGCAGGCGCAAGGGAATTTATGCTCCTCGCATTCGCGCGCCAAACCGTTGCG	1080
Db	1021	GCTGAAAAACCTGCAGGCGCAAGGGAATTTATGCTCCTCGCATTCGCGCGCCAAACCGTTGCG	1080
Qy	1081	CCGGGTGAAAGCCGGATTTCGAAGCTTCGGGCAGCAGGTCGAGATCAGGGGAATGAGTTTATA	1140

1081	Db	 CCGGGTGAAGCCGGGATTCCGAAGCTTGGGCAGCAGGTCCAGATCAGGGGAATGAGTTTATA	1140
1141	Qy	 AATAAAAAAAGCACCTGAAAAAGGTGCTCTTTTGTGATGGTTTGAACCTTGTTCTTCTTCTT	1200
1141	Db	 AATAAAAAAAGCACCTGAAAAAGGTGCTCTTTTGTGATGGTTTGAACCTTGTTCTTCTTCTT	1200
1201	Qy	 ATCTTGATACATATAGAAAATAACGTCATTTTATTTTATTTTATTTTGTGCTGAAAGGTGCG	1260
1201	Db	 ATCTTGATACATATAGAAAATAACGTCATTTTATTTTATTTTATTTTGTGCTGAAAGGTGCG	1260
1261	Qy	 TTGAAGTGTGGTATGTATGTGTTTTTAAAGTATTTGAAAAACCCCTTAAAAATGGTGTGCACAG	1320
1261	Db	 TTGAAGTGTGGTATGTATGTGTTTTTAAAGTATTTGAAAAACCCCTTAAAAATGGTGTGCACAG	1320
1321	Qy	 AAAAACCCCATCTGTTTAAAGTTTATAGTCAGCTAAACAAAATACTAAATAGATGGGGGTTTT	1380
1321	Db	 AAAAACCCCATCTGTTTAAAGTTTATAGTGACTAAACAAAATACTAAATAGATGGGGGTTTT	1380
1381	Qy	 CTTTTAAATATTATGTGTCCTTAATAGTAGCATTTTATTCAGATGAAAAATCAAGGGTTTTAG	1440
1381	Db	 CTTTTAAATATTATGTGTCCTTAATAGTAGCATTTTATTCAGATGAAAAATCAAGGGTTTTAG	1440
1441	Qy	 TGGAACAAGCAAAAAAGTGAAAAAGTGAGACCAATGTGCTTAGGAAGACGAGATTATTAATAG	1500
1441	Db	 TGGAACAAGCAAAAAAGTGAAAAAGTGAGACCAATGTGCTTAGGAAGACGAGATTATTAATAG	1500
1501	Qy	 CTGAATAAGAACGGTGCTCTCCAAATATTCCTTATTTAGAAAAAGCAAACTCTAAAAATTCCT	1560
1501	Db	 CTGAATAAGAACGGTGCTCTCCAAATATTCCTTATTTAGAAAAAGCAAACTCTAAAAATTCCT	1560
1561	Qy	 GAAAAAGGAATGAGAAATAGTAGGACCAATAATAATGACTAGAGAAGAAAAAGATGAAG	1620
1561	Db	 GAAAAAGGAATGAGAAATAGTAGGACCAATAATAATGACTAGAGAAGAAAAAGATGAAG	1620
1621	Qy	 ATTGTTTCATGAATTAAGGAAACGAATATTGGATAAATATGGGGATGATGTTTAAGGCTATT	1680
1621	Db	 ATTGTTTCATGAATTAAGGAAACGAATATTGGATAAATATGGGGATGATGTTTAAGGCTATT	1680
1681	Qy	 GGTGTTTTATGGCTCTCTTGCTCGTCAGACTGATGGGCCCTATTCCGATATTGAGATGATG	1740
1681	Db	 GGTGTTTTATGGCTCTCTTGCTCGTCAGACTGATGGGCCCTATTCCGATATTGAGATGATG	1740
1741	Qy	 TGTTGTCATGTC AACAGAGGAAGCAGAGTTTCAGCCATGAAATGACAACCCGGTCAGTGGAAAG	1800
1741	Db	 TGTTGTCATGTC AACAGAGGAAGCAGAGTTTCAGCCATGAAATGACAACCCGGTCAGTGGAAAG	1800
1801	Qy	 GTGGAAGTGAAATTTTGATAGCGAAGAGATTTCTACTAGATTATGCAATCTCAGGTGGAAATCA	1860
1801	Db	 GTGGAAGTGAAATTTTGATAGCGAAGAGATTTCTACTAGATTATGCAATCTCAGGTGGAAATCA	1860
1861	Qy	 GATTTGGCCGCTTACATCGTCAATTTTTTCTCTATTTTGCGCATTTATGATTTTCAGGTGGA	1920
1861	Db	 GATTTGGCCGCTTACATCGTCAATTTTTTCTCTATTTTGCCGATTTATGATTCAGGTGGA	1920
1921	Qy	 TACTTAGAGAAAGTGATCAAACTGCTAAATCGGTAGAGGCCAACCGTTCACAGATGCG	1980
1921	Db	 TACTTAGAGAAAGTGATCAAACTGCTAAATCGGTAGAGGCCAACCGTTCACAGATGCG	1980
1981	Qy	 ATTTTGGCCCTTATCGTAGAAGAGCTGTTTGAATATGCAAGGCAAAATGGCGTAAATATTCGT	2040
1981	Db	 ATTTTGGCCCTTATCGTAGAAGAGCTGTTTGAATATGCAAGGCAAAATGGCGTAAATATTCGT	2040
2041	Qy	 GTGCAGGACCGACAACATTTCTACCATCTTGACTGTACAGGTAGCAATGCGAGTGCC	2100
2041	Db	 GTGCAGGACCGACAACATTTCTACCATCTTGACTGTACAGGTAGCAATGCGAGTGCC	2100
2101	Qy	 ATGTTGATTTGGTCTGCATCATCGCATCTGTTTATACGACGAGCGCTTCGGTCTTTAACTGAA	2160
2101	Db	 ATGTTGATTTGGTCTGCATCATCGCATCTGTTTATACGACGAGCGCTTCGGTCTTTAACTGAA	2160
2161	Qy	 GCAGTTTAAGCAATCAGATCTTCTTCAAGGTTATGACCAATCTGTGCGAGTTTCGTAAATGCT	2220

Db 2161 GCAGTTAAGCAATCAGATCTTCTCAGGTTATGACCAATCTGTGCCAGTTCTGTAATGTCT 2220  
QY 2221 GGTCAACTTTCCGACTCTGAGAACTTCTGAACTCGCTAGAGAAATTTCTGGAATGGATT 2280  
Db 2221 GGTCAACTTTCCGACTCTGAGAACTTCTGAACTCGCTAGAGAAATTTCTGGAATGGATT 2280  
QY 2281 CAGAGTGGACAGAAACGACACGATATATAGTGGATGTGCAAAACGACATACCATTTTGA 2340  
Db 2281 CAGAGTGGACAGAAACGACACGATATATAGTGGATGTGCAAAACGACATACCATTTTGA 2340  
QY 2341 ATTGGAAGCGCGATTTGAGTCTTACGGATGTGTAATAGGAACGCTCTTTGAGGCG 2400  
Db 2341 ATTGGAAGCGCGATTTGAGTCTTACGGATGTGTAATAGGAACGCTCTTTGAGGCG 2400  
QY 2401 CGAAGCGGCGCACGATCTCAATATCGGCACATATTTGATCGTGGCAAGCGGACAGGTC 2460  
Db 2401 CGAAGCGGCGCACGATCTCAATATCGGCACATATTTGATCGTGGCAAGCGGACAGGTC 2460  
QY 2461 CGTCTAACAGAGAAGTGATCGATCGTATAGTGGTTCAGGAAATTTAAAGACGATG 2520  
Db 2461 CGTCTAACAGAGAAGTGATCGATCGTATAGTGGTTCAGGAAATTTAAAGACGATG 2520  
QY 2521 GACTGAAGATTTTGTGATGCTTGGACTGTGTAAGCCAGACAGGCGGACGCTCAAAG 2580  
Db 2521 GACTGAAGATTTTGTGATGCTTGGACTGTGTAAGCCAGACAGGCGGACGCTCAAAG 2580  
QY 2581 ATCAGAGGTAGACCGCTAATCAATTAATTTGATATCGTCACAGAGAAACCATTTCAAACA 2640  
Db 2581 ATCAGAGGTAGACCGCTAATCAATTAATTTGATATCGTCACAGAGAAACCATTTCAAACA 2640  
QY 2641 TCACAACTCACATACATACGATGACAGAGTCAATCGGTTGAAATCGCAAAAGAAATCGG 2700  
Db 2641 TCACAACTCACATACATACGATGACAGAGTCAATCGGTTGAAATCGCAAAAGAAATCGG 2700  
QY 2701 GGCTGTCTCGTGTTCAGGCGCCATTTATCGGATGAAGGAGACGAAACAGGATGTCTATTG 2760  
Db 2701 GGCTGTCTCGTGTTCAGGCGCCATTTATCGGATGAAGGAGACGAAACAGGATGTCTATTG 2760  
QY 2761 ACATCGCAAAAGCTTGAAGGCTTTCAGCGGATTCATTTCTGTGAAATTTTGTGATG 2820  
Db 2761 ACATCGCAAAAGCTTGAAGGCTTTCAGCGGATTCATTTCTGTGAAATTTTGTGATG 2820  
QY 2821 CAATTGATGACGCGCTTGAAGGCTTTCAGCGGATTCATTTCTGTGAAATTTTGTGATG 2880  
Db 2821 CAATTGATGACGCGCTTGAAGGCTTTCAGCGGATTCATTTCTGTGAAATTTTGTGATG 2880  
QY 2881 TGCTGGCGCTGTTCCGTTTATCAATCCATCAAAAGAAATTCGATTTCCGGAGGAGAG 2940  
Db 2881 TGCTGGCGCTGTTCCGTTTATCAATCCATCAAAAGAAATTCGATTTCCGGAGGAGAG 2940  
QY 2941 AGGTCAATCTCCGACATTCGACGCTTACGGCTTTTACGCGCAAACTCCATTTTGTGCG 3000  
Db 2941 AGGTCAATCTCCGACATTCGACGCTTACGGCTTTTACGCGCGCAAACTCCATTTTGTGCG 3000  
QY 3001 GAGACTACTTTAACTTACGCGGCAAGAGGAGAGGAGGATCATATAATCTCGAGTGAAT 3060  
Db 3001 GAGACTACTTTAACTTACGCGGCAAGAGGAGAGGAGGATCATATAATCTCGAGTGAAT 3060  
QY 3061 TAGGCTTTGAAGTTGAATCAGTCAAGAAATGAAGGCTAGTTAAAGTGGGAAAGCTGAA 3120  
Db 3061 TAGGCTTTGAAGTTGAATCAGTCAAGAAATGAAGGCTAGTTAAAGTGGGAAAGCTGAA 3120  
QY 3121 AGAATCAATAAAGCAATCGGTATGATGTCGAATTC 3156  
Db 3121 AGAATCAATAAAGCAATCGGTATGATGTCGAATTC 3156

## RESULT 2

AR435692 LOCUS 3156 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 1 from patent US 6656721.  
AR435692  
AR435692.1 GI:40198766

## KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 3156)

AUTHORS Hohmann,H.-P., Mouncey,N.J., Schlieker,H.W. and Stebbins,J.W.

TITLE Polynucleotide portions of the biotin operon from B. subtilis for

JOURNAL use in enhanced fermentation

Patent: US 6656721-A 1 02-DEC-2003;

Roche Vitamins, Inc.; Parsippany, NJ

FEATURES

source Location/Qualifiers

1..3156

/organism="unknown"

/mol\_type="genomic DNA"

## ORIGIN

Query Match 100.0%; Score 3156; DB 6; Length 3156;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGATCCACGAGGTTACGAGCCTTGAAGATTGATTCCTGGTTAAACGAGCGGTTAGACAGA 60

Db 1 GGATCCACGAGGTTACGAGCCTTGAAGATTGATTCCTGGTTAAACGAGCGGTTAGACAGA 60

QY 61 ATGAAAGAACGCGGCTACATCGTAACTCGGTCAATGGATGGAGCGCGGTTCCAGAG 120

Db 61 ATGAAAGAACGCGGCTACATCGTAACTCGGTCAATGGATGGAGCGCGGTTCCAGAG 120

QY 121 AGGAATATTGATGGCGAAATCAACGGTCTGCTCTCAACAAATTTATTTAGGCTCGCA 180

Db 121 AGGAATATTGATGGCGAAATCAACGGTCTGCTCTCAACAAATTTATTTAGGCTCGCA 180

QY 181 AGCGATAGACGTTTGTATCGATGAGCCCAACAGCAATTTGGGACAGGAAGC 240

Db 181 AGCGATAGACGTTTGTATCGATGAGCCCAACAGCAATTTGGGACAGGAAGC 240

QY 241 AGCGGTTCACTTTTAAACGACAGCAATCGGTCTGCATGAAAGCTTAGAAAAGAAAT 300

Db 241 AGCGGTTCACTTTTAAACGACAGCAATCGGTCTGCATGAAAGCTTAGAAAAGAAAT 300

QY 301 GCCAGCTTTAACTGACAGAGCGGCGCTGCTTTTCGAGCGGTTACTTTGGCCAAATGTC 360

Db 301 GCCAGCTTTAACTGACAGAGCGGCGCTGCTTTTCGAGCGGTTACTTTGGCCAAATGTC 360

QY 361 GGTGTCCTTTTCACTTTTAAACGAGGAGATGTCATTTTAAAGTGACAGCTCAATCAT 420

Db 361 GGTGTCCTTTTCACTTTTAAACGAGGAGATGTCATTTTAAAGTGACAGCTCAATCAT 420

QY 421 GCAAGTATGATCGACGCGCTGCGACCTTTCTAAGGCTGATACAGTTGTTATCGGCATATT 480

Db 421 GCAAGTATGATCGACGCGCTGCGACCTTTCTAAGGCTGATACAGTTGTTATCGGCATATT 480

QY 481 GATATGAATGATCTTGAACAAAGCTGAAATGAACACAGCGTTATCAGCGCGGTTTATC 540

Db 481 GATATGAATGATCTTGAACAAAGCTGAAATGAACACAGCGTTATCAGCGCGGTTTATC 540

QY 541 GTAACAGAGGAGTATTTCAGCATGGATGGCACAATCGCCCTCTTGTATCAGATCATCTCA 600

Db 541 GTAACAGAGGAGTATTTCAGCATGGATGGCACAATCGCCCTCTTGTATCAGATCATCTCA 600

QY 601 CTTTGGAAACGCTTATCATGCTTTGCTGCTGTTGATGATGCCCAACGAAAGGTTTTG 660

Db 601 CTTTGGAAACGCTTATCATGCTTTGCTGCTGTTGATGATGCCCAACGAAAGGTTTTG 660

QY 661 GCGGATTCGGGACAGAGGAGTGAATCTTTGGTGTGTTGTCGCGACATTTGTTATCGGC 720

Db 661 GCGGATTCGGGACAGAGGAGTGAATCTTTGGTGTGTTGTCGCGACATTTGTTATCGGC 720

QY 721 ACCTTAAGCAAAAGCTGTGCGGGAAGGAGTTTTTGGCGAGATCAGCGGCTTCTCATC 780

Db 721 ACCTTAAGCAAAAGCTGTGCGGGAAGGAGTTTTTGGCGAGATCAGCGGCTTCTCATC 780

QY 781 GACTTTTGTGTAACCATGCCAGAACATTTATCTTTTCAACCGCTATTTCCGCCAGCCAGC 840

781	Db	 GAC T T T T T G C T G T G A A C C A T G C C A G A A C A T T A T C T T T C A A C C G C T A T T C G C C A G C C A G C	840
841	Qy	T G T C G C G C T G T C A C A G G C T T C A A C A T C A T T G A G C C A G C A G G A A A A A C C A C A G C T T	900
841	Db	T G T C G C G C T G T C A C A G G C T T C A A C A T C A T T G A G C C A G C A G G A A A A A C C A C A G C T T	900
901	Qy	T T A T T T T C T T A T A T C A G C A T G A T C A G A A C C A G C T G A A G A A T A T G G T A T G T G T G A A	960
901	Db	T T A T T T T C T T A T A T C A G C A T G A T C A G A A C C A G C T G A A G A A T A T G G T A T G T G T G A A	960
961	Qy	G G A G A T C A C A C C G A T T T C T G T A G T C A T T G G C G A T G G C G A T G G C T C T A T T T	1020
961	Db	G G A G A T C A C A C C G A T T T C T G T A G T C A T T G G C G A T G G C G A T G G C T C T A T T T	1020
1021	Qy	G C T C A A A A C T G C A G G C C A A G G A T T A T G C T C C T G C C A T C G G C G C C A A C C G T T G G	1080
1021	Db	G C T G A A A A C T G C A G G C C A A G G A T T A T G C T C C G C C A T C G G C G C C A A C C G T T G G	1080
1081	Qy	C C G G T G A A A C C G G A T T C G A A G C T T C G G C A G C A G G T C G A G A T C A G G A T G A G T T A T A	1140
1081	Db	C C G G T G A A G C C G G A T T C G A G C T T C G G C A G C A G G T C G A G A T C A G G A T G A G T T A T A	1140
1141	Qy	A A A T A A A A A A A G C A C T G A A A G G T G C T T T T T G A T G T T T T G A A C T T G T T C T T C T T	1200
1141	Db	A A A T A A A A A A A G C A C T G A A A G G T G C T T T T T G A T G T T T T G A A C T T G T T C T T C T T	1200
1201	Qy	A T C T T G A T A C A T A T A G A A A T A A C G T C A T T T T T A T	1260
1201	Db	A T C T T G A T A C A T A T A G A A A T A A C G T C A T T T T T A T	1260
1261	Qy	T T G A G T T G G T A T G T G T	1320
1261	Db	T T G A G T T G G T A T G T G T	1320
1321	Qy	A A A A C C C C A T C T G T T A A G T T A A G T A C T A A A C A A A T A A C A A A T A A A T A G A T C G G G T T	1380
1321	Db	A A A A C C C C A T C T G T T A A G T T A A G T A C T A A A C A A A T A A C A A A T A A A T A G A T C G G G T T	1380
1381	Qy	C T T T T A T A T T G T G T C C T A A T A G T A G C A T T T T T C A G A T G A A A A A A A A A A A C A A G G T T T A G	1440
1381	Db	C T T T T A T A T T G T G T C C T A A T A G T A G C A T T T T T C A G A T G A A A A A A A A A A A C A A G G T T T A G	1440
1441	Qy	T G G A C A G A C A A A A G T G G A A A G T G A G A C C A T G C T T A G G A G A G C A G T A T T A A T A G	1500
1441	Db	T G G A C A G A C A A A A G T G G A A A G T G A G A C C A T G C T T A G G A G A G C A G T A T T A A T A G	1500
1501	Qy	C T G A A T A A G A C G T G C T C C A A A T A T C T T T T T A G A A A G C A A A T C T A A A A T A T C T	1560
1501	Db	C T G A A T A A G A C G T G C T C C A A A T A T C T T T T T A G A A A G C A A A T C T A A A A T A T C T	1560
1561	Qy	G A A A G G G A T G A A A T A G T G A T G G A C C A A T A A T A T G A T G A T G A G A G A A A A A G A T G A G	1620
1561	Db	G A A A G G G A T G A A A T A G T G A T G G A C C A A T A A T A T G A T G A T G A G A G A A A A A G A T G A G	1620
1621	Qy	A T T C T C A T G A A T T A A G G A C G A A T T G G A A A T T G G A A A T T G G G A T G A T T T A A G G C T A T T	1680
1621	Db	A T T C T C A T G A A T T A A G G A C G A A T T G G A A A T T G G A A A T T G G G A T G A T T T A A G G C T A T T	1680
1681	Qy	G G T G T T T A T G G C T C T T G G T C G T C A G A T G G G C C T A T T C G G A T T T G G A T T G A G A T G	1740
1681	Db	G G T G T T T A T G G C T C T T G G T C G T C A G A T G G G C C T A T T C G G A T T T G G A T T G A G A T G	1740
1741	Qy	T G T C A T G T C A C A G A G A G C A G A G T T C A G C C A T G A T G G A T G G A C A C C G G T G A G T G G A G	1800
1741	Db	T G T C A T G T C A C A G A G A G C A G A G T T C A G C C A T G A T G G A T G G A C A C C G G T G A G T G G A G	1800
1801	Qy	G T G A G T G A A T T T T G A T G C G A G A G A T T C A T A G A T T A T G C A T C T C A G T G G A T C A	1860
1801	Db	G T G A G T G A A T T T T G A T G C G A G A G A T T C A T A G A T T A T G C A T C T C A G T G G A T C A	1860
1861	Qy	G A T T G G C G C T T A C A C A T G G T C A A T T T T T C T A T T T T G C C G A T T A T G A T T C A G G T G G A	1920

Db	1861	GATTGGCCGCTTACACATGGTCAATTTTCTCTATTTTGGCCGATTTATGATTCAGGTGGA	1921
Qy	1921	TACTTAGAGAAAGTGATCAAACTGCTAAATCGGTAGAGAGCCAAACGTTTCCACGATGCG	1980
Db	1921	TACTTAGAGAAAGTGATCAAACTGCTAAATCGGTAGAGAGCCAAACGTTTCCACGATGCG	1980
Qy	1981	ATTGTGCCCCTTATCGTAGAAGAGCTGTTTGAATATGCAAGGCCAAATGGCGGTAAATTCGT	2040
Db	1981	ATTGTGCCCCTTATCGTAGAAGAGCTGTTTGAATATGCAAGGCCAAATGGCGGTAAATTCGT	2040
Qy	2041	GTGCAAGGACCGACAACTTCTACCATCCTTGACCTGTACAGTAGCAATGGCAGGTGCC	2100
Db	2041	GTGCAAGGACCGACAACTTCTACCATCCTTGACCTGTACAGTAGCAATGGCAGGTGCC	2100
Qy	2101	ATGTTGATTCGTCGTGCATCATCGCATCTGTTATACGACGAGCGCTTCGGTCTTAACTGAA	2160
Db	2101	ATGTTGATTCGTCGTGCATCATCGCATCTGTTATACGACGAGCGCTTCGGTCTTAACTGAA	2160
Qy	2161	GCAGTTAAGCAATCAGATCTTCTTCAGGTATATGACCATCTCTGFCGCCAGTTCGTAATGTCT	2220
Db	2161	GCAGTTAAGCAATCAGATCTTCTTCAGGTATATGACCATCTCTGFCGCCAGTTCGTAATGTCT	2220
Qy	2221	GGTCAACTTTCCGACTCTGAGAAACTTCTGGAACTCGCTAGAGAAATTTCTGGAATGGGATT	2280
Db	2221	GGTCAACTTTCCGACTCTGAGAAACTTCTGGAACTCGCTAGAGAAATTTCTGGAATGGGATT	2280
Qy	2281	CAGGAGTGGACAGAAAGACACGCGATATATAGTGGATGTGTCAAAAAGCATACCATTTTGA	2340
Db	2281	CAGGAGTGGACAGAAAGACACGCGATATATAGTGGATGTGTCAAAAAGCATACCATTTTGA	2340
Qy	2341	ATTTCGAAAGCGCCGATTTGAGTCTTACCGGATGTGTAATAAGGAAACGCTGCTTGAAGCGG	2400
Db	2341	ATTTCGAAAGCGCCGATTTGAGTCTTACCGGATGTGTAATAAGGAAACGCTGCTTGAAGCGG	2400
Qy	2401	CGAAGCGGGCGACGATCTGAATATACGGACATATTTGTATCTGTGGCAAGCGGCAGAGTTC	2460
Db	2401	CGAAGCGGGCGACGATCTGAATATACGGACATATTTGTATCTGTGGCAAGCGGCAGAGTTC	2460
Qy	2461	CGTCTAACAGAGAAGTGGATCGAGTCTGAGATCGCGTTTCAAGGAAATTTAAAGACACGTATG	2520
Db	2461	CGTCTAACAGAGAAGTGGATCGAGTCTGAGATCGCGTTTCAAGGAAATTTAAAGACACGTATG	2520
Qy	2521	GACTGAAAGATTTGTGCAATGTCTTTGGAATGTTGAAGCCAGACGAGCGGAAAGCGCTCAAAG	2580
Db	2521	GACTGAAAGATTTGTGCAATGTCTTTGGAATGTTGAAGCCAGACGAGCGGCTCAAAG	2580
Qy	2581	ATGCAGAGGTAGACCGGCTATTAATCATTAATTTGAATACGTCTACAGAGAAACCATTCACAA	2640
Db	2581	ATGCAGAGGTAGACCGGCTATTAATCATTAATTTGAATACGTCTACAGAGAAACCATTCACAA	2640
Qy	2641	TCACAACTCCACATACATACGATGACAGAGTCAATACGGTTGAAATTCGCAAAAAGAAATCGG	2700
Db	2641	TCACAACTCCACATACATACGATGACAGAGTCAATACGGTTGAAATTCGCAAAAAGAAATCGG	2700
Qy	2701	GGCTGTCTCCGTGTTTCAGGCGCCATTTATCGGGAATGAAGGAGACGAAACAGGATGTCAATG	2760
Db	2701	GGCTGTCTCCGTGTTTCAGGCGCCATTTATCGGGAATGAAGGAGACGAAACAGGATGTCAATG	2760
Qy	2761	ACATCGCCAAAAGCTTGAAGGCTCTTGAACGGGATTCATCTCTGTGGAATTTTTCGATG	2820
Db	2761	ACATCGCCAAAAGCTTGAAGGCTCTTGAACGGGATTCATCTCTGTGGAATTTTTCGATG	2820
Qy	2821	CAATTGATGACCGCGTTAGAGGGCGTCAACCAATTTAAACCCGCTGTATTGTTTAAAG	2880
Db	2821	CAATTGATGACCGCGTTAGAGGGCGTCAACCAATTTAAACCCGCTGTATTGTTTAAAG	2880
Qy	2881	TGCTGGCGCTGTTCCGTTTATCAATCCATCAAAGAAATTCGCAATTCGGAGGAAGAG	2940
Db	2881	TGCTGGCGCTGTTCCGTTTATCAATCCATCAAAGAAATTCGCAATTCGGAGGAAGAG	2940
Qy	2941	AGGTCAATCTCCGACATTCGACGCAATTAGGCTTTTACGCGCGAAACTCAATTTTGTGTCG	3000
Db	2941	AGGTCAATCTCCGACATTCGACGCAATTAGGCTTTTACGCGCGAAACTCAATTTTGTGTCG	3000





Db 1561 GAAAGGAATGAGAAATAGTAGTGAATGAGCAATAAATAGCTAGAGAAAGAAATGAAG 1620  
Qy 1621 ATTGTTCAATGAAATTAAGGAACGAATATTGATGAAATATGCGGATGATGTTAAAGCCTATT 1680  
Db 1621 ATTGTTCAATGAAATTAAGGAACGAATATTGATGAAATATGCGGATGATGTTAAAGCCTATT 1680  
Qy 1681 GGTGTTTATGCTCTCTTGCTCGTCAGACTCATGGGCCCTATTCCGATATTGAGATGATG 1740  
Db 1681 GGTGTTTATGCTCTCTTGCTCGTCAGACTCATGGGCCCTATTCCGATATTGAGATGATG 1740  
Qy 1741 TGTGTCATGTCACAGAGGAAGAGAGTTTCAGCCATGAATGGAACAACCGGTGAGTGAAG 1800  
Db 1741 TGTGTCATGTCACAGAGGAAGAGAGTTTCAGCCATGAATGGAACAACCGGTGAGTGAAG 1800  
Qy 1801 GTGGAATGAAATTTGATAGCGAAGAGATTTCTACTAGATTATGATCTCAGGTGGAATCA 1860  
Db 1801 GTGGAATGAAATTTGATAGCGAAGAGATTTCTACTAGATTATGATCTCAGGTGGAATCA 1860  
Qy 1861 GATTGGCCGCTTACACATGGTCAATTTTCTATTTTCCGATTTATGATTTCAGGTGGA 1920  
Db 1861 GATTGGCCGCTTACACATGGTCAATTTTCTATTTTCCGATTTATGATTTCAGGTGGA 1920  
Qy 1921 TACTTAGAGAAAGTGTATCAAACTGCTAAATCGGTAGAGCCCAAAAGCTTCCACGATGCG 1980  
Db 1921 TACTTAGAGAAAGTGTATCAAACTGCTAAATCGGTAGAGCCCAAAAGCTTCCACGATGCG 1980  
Qy 1981 ATTTGTGCCCTTATCGTAGAAGAGTGTTCGAATATGCGAGCAAAATGGCGTAATTCGT 2040  
Db 1981 ATTTGTGCCCTTATCGTAGAAGAGTGTTCGAATATGCGAGCAAAATGGCGTAATTCGT 2040  
Qy 2041 GTGCAAGGACCGCAACAATTTCTACCATCTTGACTGTACAGGTAGCAATGGCAGGTGCC 2100  
Db 2041 GTGCAAGGACCGCAACAATTTCTACCATCTTGACTGTACAGGTAGCAATGGCAGGTGCC 2100  
Qy 2101 ATGTTGATTGCTGTCATCGCATCTGTAATATACGACGAGCGCTTCGGTCTTAACTGAA 2160  
Db 2101 ATGTTGATTGCTGTCATCGCATCTGTAATATACGACGAGCGCTTCGGTCTTAACTGAA 2160  
Qy 2161 GCAGTTAAGCAATCAGATCTTCCCTCAGGTTATGACCATCTGCGCAGTTTCGTAATGCT 2220  
Db 2161 GCAGTTAAGCAATCAGATCTTCCCTCAGGTTATGACCATCTGCGCAGTTTCGTAATGCT 2220  
Qy 2221 GGTCAACTTTCCGACTCTGAGAACTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATT 2280  
Db 2221 GGTCAACTTTCCGACTCTGAGAACTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATT 2280  
Qy 2281 CAGGAGTGGACAGAAACGACACGGATATATAGTGGATGTGTCAAAACGCAATACCAATTTGA 2340  
Db 2281 CAGGAGTGGACAGAAACGACACGGATATATAGTGGATGTGTCAAAACGCAATACCAATTTGA 2340  
Qy 2341 ATTCGAAGCCCGATTGAGTCTTACCGGATGTTGAATAGGAACCGTCTGTAAGGCG 2400  
Db 2341 ATTCGAAGCCCGATTGAGTCTTACCGGATGTTGAATAGGAACCGTCTGTAAGGCG 2400  
Qy 2401 CGAAGCGGCGCACGATCTGAAATATCGGCACATATTGTATCGTGGCAAGCGCGCAGAGGTC 2460  
Db 2401 CGAAGCGGCGCACGATCTGAAATATCGGCACATATTGTATCGTGGCAAGCGCGCAGAGGTC 2460  
Qy 2461 CGTCTAACAGAGAAAGTGGATCAGTCTGATGATGCGGTTTCAGGAAATTAAGAGACGTAATG 2520  
Db 2461 CGTCTAACAGAGAAAGTGGATCAGTCTGATGATGCGGTTTCAGGAAATTAAGAGACGTAATG 2520  
Qy 2521 GACTGGAATTTGTGCATGCTTGGACGTTTGAAGCCAGACAGCGGAGCGGCTCAAG 2580  
Db 2521 GACTGGAATTTGTGCATGCTTGGACGTTTGAAGCCAGACAGCGGAGCGGCTCAAG 2580  
Qy 2581 ATGCAGGAGTAGACCGCTTATATATATTTGAAATACGTCACAGAGAAACCATTTCAACA 2640  
Db 2581 ATGCAGGAGTAGACCGCTTATATATATTTGAAATACGTCACAGAGAAACCATTTCAACA 2640  
Qy 2641 TCACAACTTCACATACGATGACAGAGTCAATACGGTTGAAATCGCAAAAAGAAATCGG 2700  
Db 2641 TCACAACTTCACATACGATGACAGAGTCAATACGGTTGAAATCGCAAAAAGAAATCGG 2700

Qy 2701 GGCTGTCTCGTGTTCAGGCGCCATTATCGGATGAGGAGACGAAACAGGATGTCATTG 2760  
Db 2701 GGCTGTCTCGTGTTCAGGCGCCATTATCGGATGAGGAGACGAAACAGGATGTCATTG 2760  
Qy 2761 ACATCGCAAAAGGTTGAAGGCTCTTGACGCGGATTCATTTCCTGTGAATTTTTTGCATG 2820  
Db 2761 ACATCGCAAAAGGTTGAAGGCTCTTGACGCGGATTCATTTCCTGTGAATTTTTTGCATG 2820  
Qy 2821 CAATTGATGCGACCGCTTAGAAGCGTCAACGAATTAACCCGCTGATTTTAAAG 2880  
Db 2821 CAATTGATGCGACCGCTTAGAAGCGTCAACGAATTAACCCGCTGATTTTAAAG 2880  
Qy 2881 TGCTGGCGCTGTTCCGTTTATCAATCCATCAAAAGAAATTCGCAATTTCCGAGGAAAGAG 2940  
Db 2881 TGCTGGCGCTGTTCCGTTTATCAATCCATCAAAAGAAATTCGCAATTTCCGAGGAAAGAG 2940  
Qy 2941 AGGTCAATCTCCGCACATTCGACGCAATTAGGGCTTTACGCCGCAAACTCCATTTTTGTCG 3000  
Db 2941 AGGTCAATCTCCGCACATTCGACGCAATTAGGGCTTTACGCCGCAAACTCCATTTTTGTCG 3000  
Qy 3001 GAGACTACTTAAACAACTGCGCGGCAAGAGAGACGAGGATCATATAATGCTGAGTGATT 3060  
Db 3001 GAGACTACTTAAACAACTGCGCGGCAAGAGAGACGAGGATCATATAATGCTGAGTGATT 3060  
Qy 3061 TAGGCTTTGAAGTTGAATCAGTCGAAGAAATGAAGGCTAGTTTAAAGTCCGAAAAGCTGAA 3120  
Db 3061 TAGGCTTTGAAGTTGAATCAGTCGAAGAAATGAAGGCTAGTTTAAAGTCCGAAAAGCTGAA 3120  
Qy 3121 AGAATCAATAAAAGCAATCGGTATGATGTCGAATTC 3156  
Db 3121 AGAATCAATAAAAGCAATCGGTATGATGTCGAATTC 3156

RESULT 4  
AR173058  
LOCUS AR173058 8478 bp DNA linear PAT 17-DEC-2001  
DEFINITION Sequence 1 from patent US 6303377.  
ACCESSION AR173058  
VERSION AR173058.1 GI:17912549  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 8478)  
AUTHORS Bower,S.Grant., Perkins,J.B., Yocum,R.Rogers. and Pero,J.G.  
TITLE Biotin biosynthesis in Bacillus subtilis  
JOURNAL Patent: US 6303377-A 1 16-OCT-2001;  
FEATURES Location/Qualifiers  
source 1..8478  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 51.1%; Score 1612.6; DB 6; Length 8478;  
Best Local Similarity 73.0%; Pred. No. 0;  
Matches 2303; Conservative 0; Mismatches 624; Indels 226; Gaps 9;  
Qy 2 2 GATCCACGAGTTACGAGCTTGAAGATTGATTCCTGGTTAAACGAGCGGTTAGACAGAA 61  
Db 2524 GATCCACGAGTTACGAGCTTGAAGATTGATTCCTGGTTAAACGAGCGGTTAGACAGAA 2583  
Qy 62 TGAAGAGCGCGGTACATCGTAACCTCGCGTCAATGGATGGAGCCGCGTCCAGAGA 121  
Db 2584 TGAAGAGCGCGGTACATCGTAACCTCGCGTCAATGGATGGAGCCGCGTCCAGAGA 2643  
Qy 122 GGAATATTGATGGCGAAAATCAAAACGGTCTGGTCTCTCAAAACAATTTATTAGGGCTCGCAA 181  
Db 2644 GGAATATTGATGGCGAAAATCAAAACGGTCTGGTCTCTCAAAACAATTTATTAGGGCTCGCAA 2703  
Qy 182 GCGATAGACGTTGATCGATGCGAGCCCAAAACAGCAATTCGAGCAATTTGGGACAGGAAGCA 241  
Db 2704 GCGATAGACGTTGATCGATGCGAGCCCAAAACAGCAATTCGAGCAATTTGGGACAGGAAGCA 2763

QY 242 GCGTTTCAAGTTAAACGACAGGCAATTCGGTCTGGCATGAAAAGCTAGAAAAGAGATTG 301  
DB 2764 GCGTTTCAAGTTAAACGACAGGCAATTCGGTCTGGCATGAAAAGCTAGAAAAGAGATTG 2823  
QY 302 CAGCTTTAAACTGACAGAGCGGCCCTGCTGTTTTCGAGCGGTACTTGGCCCAATGTCG 361  
DB 2824 CAGCTTTAAACTGACAGAGCGGCCCTGCTGTTTTCGAGCGGTACTTGGCCCAATGTCG 2883  
QY 362 GTGTCCTTTCATCTCTCCAGAAAAGAGATGTCATTTTAAAGTGACAGCTCAATCATG 421  
DB 2884 GTGTCCTTTCATCTCTCCAGAAAAGAGATGTCATTTTAAAGTGACAGCTCAATCATG 2943  
QY 422 CAAGTATGATCGAGCGCTGCCGACTTTCTAAGGCTGATACAGATGTTTATCGGCATATTG 481  
DB 2944 CAAGTATGATCGAGCGCTGCCGACTTTCTAAGGCTGATACAGATGTTTATCGGCATATTG 3003  
QY 482 ATATGAATGATCTTGAACACAGCTGAATGAACACAGAGCTTATCAGCGCGCTTTATCG 541  
DB 3004 ATATGAATGATCTTGAACACAGCTGAATGAACACAGAGCTTATCAGCGCGCTTTATCG 3063  
QY 542 TAAACAGAGGATGATTCAGCATGATGCGCAATTCGCCCTCTTGATCAGATCATCTCAC 601  
DB 3064 TAAACAGAGGATGATTCAGCATGATGCGCAATTCGCCCTCTTGATCAGATCATCTCAC 3123  
QY 602 TTGCGAAACGCTATCATGCTCTCGTGTGCTGATGATGCCAGCAACAGAGATTTTGG 661  
DB 3124 TTGCGAAACGCTATCATGCTCTCGTGTGCTGATGATGCCAGCAACAGAGATTTTGG 3183  
QY 662 GCGATTCGGGACAGGAAACAGATGAATACATTTGGTGTGTTTGTCCGCAATTTATCGGCA 721  
DB 3184 GCGATTCGGGACAGGAAACAGATGAATACATTTGGTGTGTTTGTCCGCAATTTATCGGCA 3243  
QY 722 CCTTAAGCAAGCTGTTGGCGGAGAGGTTTGGCGGAGATCAGCGCTCTTCATCG 781  
DB 3244 CCTTAAGCAAGCTGTTGGCGGAGAGGTTTGGCGGAGATCAGCGCTCTTCATCG 3303  
QY 782 ACTTTTGTGTAACCATGCCAGAACATTTATCTTTCAACCGCTATTTCGCCAGCCAGCT 841  
DB 3304 ACTTTTGTGTAACCATGCCAGAACATTTATCTTTCAACCGCTATTTCGCCAGCCAGCT 3363  
QY 842 GTGCGGTGCTCAGAGGCTTTCAACATCATTTGAAGCCAGAGGAGAAAACGACAGCTTT 901  
DB 3364 GTGCGGTGCTCAGAGGCTTTCAACATCATTTGAAGCCAGAGGAGAAAACGACAGCTTT 3423  
QY 902 TATTTTCTTATATCAGCATGATCAGAACCATCTCTGAAGATATGGGTTATGCTGTAAG 961  
DB 3424 TATTTTCTTATATCAGCATGATCAGAACCATCTCTGAAGATATGGGTTATGCTGTAAG 3483  
QY 962 GAGATCACACACCGATTATTCCTGTAGTCAATTTGGCGATGCCCATATAAACCGTCTATTG 1021  
DB 3484 GAGATCACACACCGATTATTCCTGTAGTCAATTTGGCGATGCCCATATAAACCGTCTATTG 3543  
QY 1022 CTGAAAACCTGAGGCAAGGAAATTTATGCTCTGCCATTCGGCCGCCAACCGTTGCGC 1081  
DB 3544 CTGAAAACCTGAGGCAAGGAAATTTATGCTCTGCCATTCGGCCGCCAACCGTTGCGC 3603  
QY 1082 CGGCTGAAGCCGATTCGAAGCTTTGGCGAGCAGCTCGAGATCAGGGAATGAGTTTATAA 1141  
DB 3604 CGGCTGAAGCCGATTCGA-----ATTA 3627  
QY 1142 AATAAAAAAGACCTGAAAAGGTGCTTTTTTTTGTGTTTGAACCTGTTCTTTCTTA 1201  
DB 3628 CAATCAGCTCTGACACAGATGAGTATGATCATTTGCTGCAACATTTCAATCAA 3687  
QY 1202 TCTTGATACATATAGAAATACAGTCATTTTTTTATTTTATTTTATGCTGTAAGGTCGCT 1261  
DB 3688 TCGAAAAGGAGCTGCACATCATTTTGGGGGTTTTTTTGTGACGCGGAACCTGA----- 3738  
QY 1262 TGAAGTGTGATGATGTTTTTAAAGTATTCAGAAACCTTAAATTTGTTGTCACAGA 1321  
DB 3739 -----TACAGAAAGTAGGGAACCGGTTATATCCAGCGCTCTTG- 3776

QY 1322 AAAACCCCATCTGTTAAAGTTTATAAGTGACTAAACAAATACTAAATAGATGGGGTTTC 1381  
DB 3777 -----CTGCCCTTATTGAAAGACAATAATAGACATCTCGGGGTGTA 3816  
QY 1382 TTTTAATATTATGTGCTCTTAATAGTAGTACATTTTATTCAGATGAAAATCAAGGGTTTTAGT 1441  
DB 3817 TAAACCATTTTTTAAAGCGGATATCGCGCCATCATCCAGATAGTGATACAAGTTTTCTGAA 3876  
QY 1442 GGCACAGACAAAAGTGGAAAAGTGAGACCATGTGCTTAGAAGACGAGTTATTAATAGC 1501  
DB 3877 AGATATGTCGCAG----- 3890  
QY 1502 TGAATGAAGACCGGTGCTCTCAAAATATTCTTATTAGAAAAGCAAAATCTAAAATATTCTG 1561  
DB 3891 -----CCAGTCTTCTCTCTGAAGACATTAACGCTTTTTCGCTTCAA 3930  
QY 1562 AAAAGGAATGAGAAATAGTGAATGGACCAATATAATGACTAGAGAAGAAAGATGAAGA 1621  
DB 3931 GCGCCGCTTTCACCATACGTTTCAGGGGAAACCTTTGAGGGAAAGACTGTCACCATGGGAAGA 3990  
QY 1622 TTGTTTATGAAATTTAAAGGAACGAAATATTGGATAAATAATGGGGATGATGTTTAAAGCTATTG 1681  
DB 3991 GGTTTTAAGCCATTTGGGGCGGATTAGAGAAA-----CATGAATGCTTCATCGTAG 4043  
QY 1682 GTGTTTATGGCTCTCTTTGGTGTGTCAGATGATGGCCCTTATTCGGATATTAGATGATGAT 1741  
DB 4044 AAGGTGCGAGCGGTATTTCTGTGCCATTTGGGAGAGGACTATTTGGTCAGTCATGTCATAA 4103  
QY 1742 GTGTCATGTCAACAGAGGAGCAGAGTTTCAGCCATGNAATGGACAAACCGGTGAGTGAAGG 1801  
DB 4104 -----AAGCGTTGCAGCTTCCCATGATTTATTTGGCGGCTCTTCGCCCTTGGAAAC 4153  
QY 1802 TGGAAAGTGAATTTTGTATAGCGAAGAGATTCTACTAGATTATGCTCTCAGGTGGAATCAG 1861  
DB 4154 ATTAATCATACCTTTTAACTGTCAANTATGAGAAAGCATGGGCTTCCCAATCGCCGGA 4213  
QY 1862 ATTGGCGGCTTACATGCTCAATTTTTCTCTATTTTGGCGATTTATGATTCAGGTGAT 1921  
DB 4214 ATTTATCATCAATGGAATCAGTGACTCTCTGTGATGAAGATGAAAACCAATCTCTGAGATG 4273  
QY 1922 ACTTAGAAGAGTATCAAACTGCTAAATCGGTAGAGCCCAACCGTCCACGATCGGA 1981  
DB 4274 ATTGAGCGCTTATTCGGGTGTCGGAATTTAGGGGTTCGCCAAAGCTTGCACAGTG- 4330  
QY 1982 TTTGTGCCCTTATCGTAGAAGAGCTGTTTGAATATGACAGGCAAAATGGCGTAATATTCTGT 2041  
DB 4331 -----ACGAAGAAACGGTCTCATATGTTGAAGACCAT- 4367  
QY 2042 TGCAGGACCGACAACATTTCTACCATCTTGTACTGTACAGGTAGCAATGGCAGGTGCCA 2101  
DB 4368 -----TCAATCTATCATTTACTGATGAATCAAGTGGGGTATGA 4405  
QY 2102 TGTGATTGCTCTGCATCATCGCATCTGTTATACGAGCGGCTTCGGTCTTAACTCAAG 2161  
DB 4406 GAATGAATCAATGATGGAATCTCGAGACCGGGTGTGGCTGGAGCAAGATGATGACG 4465  
QY 2162 CAGTTAAGCAATCAGATCTTCTTTCAGGTTATGACCATCTGTGCGAGTTCGTATGCTG 2221  
DB 4466 AAGAGGGCTTTCAATATTTACATTTGCTGTGATGAAGATATTTTGTCTATTATGACCGGG 4525  
QY 2222 GTCAACTTTCCGACTCTGAGAAAACCTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTC 2281  
DB 4526 CTTTTTACATCAGAAAAACATTTTACGAAAAAAGGATTAAGCTCAATATGATGATGATG 4585  
QY 2282 AGGAGTGGACAGAACGACACGATATATAGTGGATGTGTCAAAACGCAATACCATTTTGA 2341  
DB 4586 CGAAATCCGGGCTCTGCGCCGGAACCTGCGGCTATTGTTTCA-----GTCTCGCAT 4637  
QY 2342 TTCGAAAGCGCCGATTTAGTCTTTACCGGATGTTGAATAAGGAAACGCTGCTTGAAGCGC 2401  
DB 4638 TTCGAAAGCGCCGATTTAGTCTTTACCGGATGTTGAATAAGGAAACGCTGCTTGAAGCGC 4697  
QY 2402 GAAGCGGGCGCACGATCTGAATATTCGGCAATATTGTTATCGTGGCAAGCGGACAGGTCC 2461

Db	4698	GAAGCGCGCACGATCTGAATATATCGGCACATATTGTATGTCGCAAGCGCGAGGTCC	4757
Qy	2462	GTCTAAACAGAGAGTGGATCAGGTGATAGTCGGTTTCAGGAAATTAAGAGAGATATGG	2521
Db	4758	GTCTAAACAGAGAGTGGATCAGGTGATAGTCGGTTTCAGGAAATTAAGAGAGATATGG	4817
Qy	2522	ACTGAAGATTTGTGCATGCTCTTGGACTGTTGAAGCCAGAGCAGGCGAAGCGGCTCAAGA	2581
Db	4818	ACTGAAGATTTGTGCATGCTCTTGGACTGTTGAAGCCAGAGCAGGCGAAGCGGCTCAAGA	4877
Qy	2582	TGCAGGAGTAGACCGCTATAATCATAAATTTGAATACGTCACAGAGAAACCAATTCAAACAT	2641
Db	4878	TGCAGGAGTAGACCGCTATAATCATAAATTTGAATACGTCACAGAGAAACCAATTCAAACAT	4937
Qy	2642	CACAACCTCACATACATACGATGACAGAGTCAATACGGTTGAAATCGCAAAAGATCGGG	2701
Db	4938	CACAACCTCACATACATACGATGACAGAGTCAATACGGTTGAAATCGCAAAAGATCGGG	4997
Qy	2702	GCTGTCTCCGTGTTTCAGCGCGCATTTATCGGATGAAGGAGACGAAACAGGATGTCATTGA	2761
Db	4998	GCTGTCTCCGTGTTTCAGCGCGCATTTATCGGATGAAGGAGACGAAACAGGATGTCATTGA	5057
Qy	2762	CATGCCCAAAAGCTTTGAAGGCTTTGACCGGATTCCAATTCCTGTGAAATTTTTTGCATGC	2821
Db	5058	CATGCCCAAAAGCTTTGAAGGCTTTGACCGGATTCCAATTCCTGTGAAATTTTTTGCATGC	5117
Qy	2822	AATTGATGGCAGCGCTTTAGAGCGGTCAACGGAATTAACCCGCTGATTTGTTAAAGT	2881
Db	5118	AATTGATGGCAGCGCTTTAGAGCGGTCAACGGAATTAACCCGCTGATTTGTTAAAGT	5177
Qy	2882	GCTGCGCTGTTTCCGTTTTTATCAATCCATCAAAAGAAATTCGCATTTCCGGAGGAAGAGA	2941
Db	5178	GCTGCGCTGTTTCCGTTTTTATCAATCCATCAAAAGAAATTCGCATTTCCGGAGGAAGAGA	5237
Qy	2942	GGTCAATCTCCGCATTCGACGCCATTAGGCGCTTTACGCCCAACTCCATTTTGTGCGG	3001
Db	5238	GGTCAATCTCCGCATTCGACGCCATTAGGCGCTTTACGCCCAACTCCATTTTGTGCGG	5297
Qy	3002	AGACTACTTAACACTGCGGGCAGAGGACGAGGATCATAAATGCTGAGTGATTT	3061
Db	5298	AGACTACTTAACACTGCGGGCAGAGGACGAGGATCATAAATGCTGAGTGATTT	5357
Qy	3062	AGGCTTTGAAGTTGAATCAGTCGAAGAAATGAAGCTAGTTTAAAGTGGAAAAAGCTGAAA	3121
Db	5358	AGGCTTTGAAGTTGAATCAGTCGAAGAAATGAAGCTAGTTTAAAGTGGAAAAAGCTGAAA	5417
Qy	3122	GAATCAATAAAGCAATCGGTATGATGCGAAT	3154
Db	5418	GAATCAATAAAGCAATCGGTATGATGCGAAT	5450

RESULT 5	AR630545	8478 bp	DNA	linear	PAT 14-FEB-2005
LOCUS	AR630545	Sequence 1 from patent US 6841366.			
DEFINITION	AR630545				
ACCESSION	AR630545.1	GI:59765691			
VERSION	Unknown.				
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unclassified.				
REFERENCE	1 (bases 1 to 8478)				
AUTHORS	Bower, S.G., Perkins, J.B., Vocum, R.R. and Pero, J.G.				
TITLE	Biotin biosynthesis in bacillus subtilis				
JOURNAL	Patent: US 6841366-A 1 11-JAN-2005;				
	DSM IP Assets B.V.; Heerlen;				
	NLX;				
FEATURES	Location/Qualifiers				
source	1..8478				
	/organism="unknown"				
	/mol_type="genomic DNA"				
ORIGIN					

Query Match	51.1%;	Score 1612.6;	DB 6;	Length 8478;	
Best Local Similarity	73.0%;	Pred. No. 0;			
Matches 2303;	Conservative	0;	Mismatches 624;	Indels 226;	Gaps 9;
Qy	2	GATCACGAGGTTACGAGCCTTTGAAGATTGATTCTCTGGTTAAACGAGCGGTTAGACAGAA	61		
Db	2524	GATCCACGAGGTTACGAGCCTTTGAAGATTGATTCTCTGGTTAAACGAGCGGTTAGACAGAA	2583		
Qy	62	TGAAGAAGCGGCGGTACATCGTAACTCTGGGTCAATGGATGGAGCGCGGTTCCAGAGA	121		
Db	2584	TGAAGAAGCGGCGGTACATCGTAACTCTGGGTCAATGGATGGAGCGCGGTTCCAGAGA	2643		
Qy	122	GGATATTGATGGCGAAAAATCAAAACGGTCTGTGCTCTCAAAACAAATTATTTAGGGCTCGCAA	181		
Db	2644	GGATATTGATGGCGAAAAATCAACGGTCTGTGCTCTCAAAACAAATTATTTAGGGCTCGCAA	2703		
Qy	182	GCATAGACGTTTGATCGATGAGCCCAACACAGCATTCGAGCAATTTGGGACAGGAAGCA	241		
Db	2704	GCATAGACGTTTGATCGATGAGCCCAACACAGCATTCGAGCAATTTGGGACAGGAAGCA	2763		
Qy	242	GCAGTTTACGTTTAAACGACAGCAATTCGGTCTGGCATGAAAGCTAGAAAAGAGATTG	301		
Db	2764	GCAGTTTACGTTTAAACGACAGCAATTCGGTCTGGCATGAAAGCTAGAAAAGAGATTG	2823		
Qy	302	CCAGCTTTAAACTGACAGAAGCGGCCCTGCTGTTTTTCGAGCGGTTACTTTGGCCAAATGTCG	361		
Db	2824	CCAGCTTTAAACTGACAGAAGCGGCCCTGCTGTTTTTCGAGCGGTTACTTTGGCCAAATGTCG	2883		
Qy	362	GTGTCCTTTTCACTCTTCGAGAAAAAGGAAGATGTCATTTTAAAGTGACAGCTCAATCATG	421		
Db	2884	GTGTCCTTTTCACTCTTCGAGAAAAAGGAAGATGTCATTTTAAAGTGACCAAGCTCAATCATG	2943		
Qy	422	CAAGTATCATCGAGCGGTCGCGACTTTCCTAAGGCTGATACAGTGTGTTTATCGGCATATTG	481		
Db	2944	CAAGTATGATCGAGCGGTCGCGACTTTCCTAAGGCTGATACAGTGTGTTTATCGGCATATTG	3003		
Qy	482	ATATGAATGATCTTTGAAAAACAAGCTGAATGAAACACAGCGGTTATCAGCGCGGTTTTATCG	541		
Db	3004	ATATGAATGATCTTTGAAAAACAAGCTGAATGAAACACAGCGGTTATCAGCGCGGTTTTATCG	3063		
Qy	542	TAAACAGCGGAGTATTACGATGGATGGCAAAATCGCCCTCTTGATCAGATCATCTCAC	601		
Db	3064	TAAACAGCGGAGTATTACGATGGATGGCAAAATCGCCCTCTTGATCAGATCATCTCAC	3123		
Qy	602	TTGCGAAACGGCTATCATGCTTCGTGGTCGTTTGATGATGCCCAACAGGAGTTTTCG	661		
Db	3124	TTGCGAAACGGCTATCATGCTTCGTGGTCGTTTGATGATGCCCAACAGGAGTTTTCG	3183		
Qy	662	GCGATTTCGGGACAAAGGAACGAGTGAATACCTTTGGTGTGTTTGTCCGACATTTGTTATCGGCA	721		
Db	3184	GCGATTTCGGGACAAAGGAACGAGTGAATACCTTTGGTGTGTTTGTCCGACATTTGTTATCGGCA	3243		
Qy	722	CCTTAAGCAAAAGCTTTGCGCGGAAGAGGTTTTTGGCGAGGATCAGCGGCTTTCATCG	781		
Db	3244	CCTTAAGCAAAAGCTTTGCGCGGAAGAGGTTTTTGGCGAGGATCAGCGGCTTTCATCG	3303		
Qy	782	ACTTTTTTGTCTGAACCATGCCAGAACATTTATCTTTTCAAAACCGCTATTTCGCCAGCAGCT	841		
Db	3304	ACTTTTTTGTCTGAACCATGCCAGAACATTTATCTTTTCAAAACCGCTATTTCGCCAGCAGCT	3363		
Qy	842	GTGCGGCTGCTCACGAGGCTTTTCAACATCATTTGAAGCCAGCAGGGAAGAAACGACAGCTTT	901		
Db	3364	GTGCGGCTGCTCACGAGGCTTTTCAACATCATTTGAAGCCAGCAGGGAAGAAACGACAGCTTT	3423		
Qy	902	TATTTTCTTATATCAGCATGATCAGAACCCAGTCTGAAGAAATATGGGTTATGTTGGTGAAG	961		
Db	3424	TATTTTCTTATATCAGCATGATCAGAACCCAGTCTGAAGAAATATGGGTTATGTTGGTGAAG	3483		
Qy	962	GAGATCACACACCGATTATTCCTGTAGTTCATTTGGCGATGCCCATATAAACGGTCTTATTG	1021		
Db	3484	GAGATCACACACCGATTATTCCTGTAGTTCATTTGGCGATGCCCATATAAACGGTCTTATTG	3543		



BSU51868	BSU51868	10153 bp	DNA	linear	BCT 25-OCT-1996
LOCUS	Bacillus subtilis				
DEFINITION	Bacillus subtilis biotin biosynthetic operon genes, complete and partial cds.				
ACCESSION	U51868				
VERSION	U51868.1	GI:1277024			
KEYWORDS					
SOURCE	Bacillus subtilis				
ORGANISM	Bacillus subtilis				
REFERENCE	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.				
AUTHORS	1 (bases 1 to 10000) Bower,S., Perkins,J.B., Yocum,R.R., Howitt,C.L., Rahaim,P. and Pero,J.				
TITLE	Cloning, sequencing, and characterization of the Bacillus subtilis biotin biosynthetic operon				
JOURNAL	J. Bacteriol.	178 (14),	4122-4130	(1996)	
PUBMED	8763940				
REFERENCE	2 (bases 1 to 10153)				
AUTHORS	Yocum,R.				
TITLE	Direct Submision				
JOURNAL	Submitted (20-FEB-1996) Yocum R., Omnigene Bioproducts, Inc., 763 D Concord Ave., Cambridge, MA, USA, 02139-9002				
FEATURES	Location/Qualifiers				
source	1..10153				
	/organism="Bacillus subtilis"				
	/mol_type="genomic DNA"				
gene	/db_xref="taxon:1423"				
	/map="262 degrees"				
	/clones="pBIO100, pBIO350, pBIO201"				
CDS	2119..2898				
	/gene="bioW"				
	2119..2898				
	/gene="bioW"				
	/function="biotin pathway"				
	/codon_start=1				
	/evidence=experimental				
	/transl_table=11				
	/product="pimelyl CoA synthetase"				
	/protein_id="AAB17457.1"				
	/db_xref="GI:1277025"				
	/translation="MMQETEFYSVRMSMGSHEDGGKHISGGERLIPHEMKHTYN ALLBKGLSHSGKDFPMQIQPEEYHSIKTIQPLPVHTNEVSCPEBGQKLARLLLEKE GVSDDVTEKAYEQIPEWSDVRGAVLFDIHTGKRMDDQTEKGVRSVRMDWPDANPEKWA LHSHPAHSRIKEALAKASKVSRHPAVVAELCWSDDPDYLTGYVAGKKMGVQRTAMK EYTEBCGRVFIDGSNDVNTYIHDLKQPLIIEWEDHDS"				
gene	2888..4234				
	/gene="bioA"				
CDS	2888..4234				
	/gene="bioA"				
	/function="biotin pathway"				
	/codon_start=1				
	/evidence=experimental				
	/transl_table=11				
	/product="DAPA aminotransferase"				
	/protein_id="AAB17458.1"				
	/db_xref="GI:1277026"				
	/translation="MTWDLIEKSKHLWLPFTQMKVDENPLIIESGTGIVKVDINK EYDFGSVWLVNVRHKKLELDDAIKKQLGKIAHSTLLGMNTNVPATQLAETLIDISPK KLTRVFSVSGAEAEIALKMAFYKNIKGKEQKFTAMKNGYHGTIGAVSYGSIE LFHVYGFPLMFPEKPIPYVYRSESGDPDECRDQCLRELAQLLEEHEHIEAALSIES MVQASGMI VMPGCVLAGVRELCTTYDVLMTVDBVATPGRTGKMFACHEHNVOPDLM AAGKGTGGYLPVATVATEDIYKAFYDDYENLATPFHGHSYTCNQIGCAVALENAL PESENIYQVAEKSXKLHFLQDLHALPHVGDIRQLGFMCGAELVRSKTEKYPYADR RIGYVSLMKELGMLTRPLGVDVIAFLPPLASTAEISEMVAIMKQAIHEVTSLED"				
gene	4224..5393				
	/gene="bioP"				
CDS	4224..5393				
	/gene="bioP"				
	/function="biotin pathway"				
	/standard_name="KAPA synthase"				
	/codon_start=1				
	/evidence=experimental				
	/transl_table=11				

	/product="L-alanine - pimelyl CoA ligase"				
	/protein_id="AAB17459.1"				
	/db_xref="GI:1277027"				
	/translation="MKIDSWNERLDRMKAGVHRNRSMDGAPVPERNIDGENQTVW SSNNYGLASDRRLIDAAQTALQOFGTSGSRLTTGNSVMHEKLEKXIASFUKITEAA LLFSSGYLANVGLVSLPEKEDVILSDQLNHSAMIDGCRLSKADTVVYRHHMDNDLEN KLTETORYQRRFIVTDGVFSMDGTIAPLDQIISLAKRYHAFVVDVDAHATGVLGDSGQ GTSEYFGVCPDVI VIGTLSKAVGAEAGGPAAGSAVFIDFLNHARTFIQTATPPASCAA AHEAFNIEASREKRQLLFSYISMIRTSLKMGVYVKGDHTPIIPVIGDAHKTVLFA EKLOGKGIYAPAIRPPTVAPGESRIRITITSDHSMGDIIDLHLQTFHSIGELH11"				
gene	5390..6085				
	/gene="bioD"				
CDS	5390..6085				
	/gene="bioD"				
	/function="biotin pathway"				
	/codon_start=1				
	/evidence=experimental				
	/transl_table=11				
	/product="desthbiotin synthetase"				
	/protein_id="AAB17460.1"				
	/db_xref="GI:1277028"				
	/translation="MRGFFVIGTDEVGKTVISSGLAALLKONNRHVGVYKPFSLGIS RHPDSDTSLKDMQSQTSLSHEDITPPAFKAPAPYVAGLEGKTVTMEELSHWGRI REKHECFIVGAGGISVPLGEDYLVSHVIALQLPMIIVARPLGTINTFLTIVKYAE SMGLPIAGIITINGISDSDPEDKTNPEMERLCGPVILGVTPLKLVTKETVLHMVKD HINLSLLMNQGV"				
gene	6088..7095				
	/gene="bioB"				
CDS	6088..7095				
	/gene="bioB"				
	/function="biotin pathway"				
	/codon_start=1				
	/evidence=experimental				
	/transl_table=11				
	/product="biotin synthase"				
	/protein_id="AAB17461.1"				
	/db_xref="GI:1277029"				
	/translation="MNQWMLADRVLAGAEVTDDEALSILHCPDEDILLMHGAFHIR KHYGKKVYKLNIMNAKSLGPCNCGYCSASAKPIESYRMVWNETLLEGAKRAHD LNIGTCIVASGRGPNREVDVQVDAVQEIETIKLACGLGLLKPQAKRLKADAGV DRHNNLNTSQRHSNITTSHTYDDRVNTVEIAKESGLSPCSGAIIGMKEKQDVIDI AKSLKALDADSIPVNLHAIDGTFLEGVNELNPLYCLKLVALFRFPNPSKEIRISGGR EVNLRTLQPLGLVAANSIFVGDVLTITAGQEBTEDEHKMLSDLGFEVESVEEMKASLSAK S"				
gene	7164..8351				
	/gene="bioI"				
CDS	7164..8351				
	/gene="bioI"				
	/function="pimelic acid synthesis"				
	/codon_start=1				
	/evidence=experimental				
	/transl_table=11				
	/product="cytochrome P450-like enzyme"				
	/protein_id="AAB17462.1"				
	/db_xref="GI:1277030"				
	/translation="MTIASSTASSEFLKNPVSFYDTLRAVHPYIKGSFLKYPGVVVTG YETAILLKDARFKVITPPESSSTKYQDLSHVQNMWLFQNDPHRLRLTLASGFTP RTTESQPIYIITVHLLDQVQKQKQMEVISDFAPFLASFIANIGVPBEDRQLKE WAASLQITDITFSRKALTEGNTMAVQAMAYFKELIQKRKHPQODMISMLLKRECL KLTEEEAASCTILLATAGHTVNLISNSVLCILQHPQLLKLEENPDLLIGTAYECL RYSPGNTARVASEDIDICGVTIRQEQVYVLLLGANRDPISFTNPDPFDITRSPN HLSEFGHHVYGLSSLARLEAQIAINTLLQRMFSLNADFEWRYRPLFGFRALBELPV TFS"				
gene	8428..9189				
	/gene="orf2"				
CDS	8428..9189				
	/gene="orf2"				
	/function="unknown"				
	/codon_start=1				
	/transl_table=11				
	/product="unknown, probably a redox enzyme"				
	/protein_id="AAB17463.1"				
	/db_xref="GI:1277031"				

/translation="MKKVLIAAGNGVIGRLLAEGLISDYEVTVLDKDFGPKASSIQ DAANYBELKKIPKDDAILNLAAVKIKYDINDIAEFERKMTDFYRASYLFCRAAE GOKLVPASSNHVTVYKEDKSGILLGREITTSYPLSKNLYGLVLTSEQIGHLFYLE NKLSVINLRIGTVTDEMDDLHEKERTKTLSSHDPDLSIFKAAIETNIRYGYIYAVS DNPRPWSIBSAVNELGFSQINTABELNEENGA"		gene	
9375..10153		9375..10153	
/gene="orf3"		CDS	
/genes="orf3"			
/codon_start=1			
/transl_table=11			
/product="unknown"			
/protein_id="AAB17464.1"			
/db_xref="GI:1277032"			
/translation="MGAIAHSLKNRLFDMLIYGLLMAFALICVLPFHVIAASFATVE EVVSKFPIPTFSLDAYRYIFSTDIIYKSLVSVFVTVIGTAVSMFLSSLMAYGLS RRDLIGQPLMFLVPTFPLPSGGMIPTFLNVKSLGLDLSYMALILPTAINAPNLILK NFFONTIPSSLEESAKIDGCDNLGIFFPKIVLPLSLPALATISLFFAVTYNTYMTAILY LNSDAKWPVIQVLLRQIVIVSSGMQGMSEMGSPPEPTIG"			
ORIGIN			
Query Match : 51.1%; Score 1612.6; DB 1; Length 10153;			
Best Local Similarity 73.0%; Pred. No. 0;			
Matches 2303; Conservative 0; Mismatches 624; Indels 226; Gaps 9;			
QY	2	GATCCACGAGGTTACGAGCCTTGAAAGATTGATTCCTGGTTAAACGAGCGGTTAGACAGAA	61
DB	4204	GATCCACGAGGTTACGAGCCTTGAAAGATTGATTCCTGGTTAAACGAGCGGTTAGACAGAA	4263
QY	62	TGAAAGAACGGCGGTACATCGTAAACCTGCGGTCAATFGATGAGCGCGGTTCCACAGA	121
DB	4264	TGAAAGAACGGCGGTACATCGTAAACCTGCGGTCAATFGATGAGCGCGGTTCCACAGA	4323
QY	122	GGAAATTTGATGCGGAAATCAAAACGGTCTGGTCTCTCAAAACAATATTTAGGGCTCGCAA	181
DB	4324	GGAAATTTGATGCGGAAATCAAAACGGTCTGGTCTCTCAAAACAATATTTAGGGCTCGCAA	4383
QY	182	GCGATAGACGTTTGATCGATGCGGCCCAACAGCATTTGCAGCAATTTGGGACAGGAAGCA	241
DB	4384	GCGATAGACGTTTGATCGATGCGGCCCAACAGCATTTGCAGCAATTTGGGACAGGAAGCA	4443
QY	242	GCGGTTCAAGCTTAAACGACAGGCAATTCGGTCTGGCATGAAAAGCTAGAAAAAGAAATTG	301
DB	4444	GCGGTTCAAGCTTAAACGACAGGCAATTCGGTCTGGCATGAAAAGCTAGAAAAAGAAATTG	4503
QY	302	CCAGCTTTAAACTGACAGAGCGGCCCTGCTGTTTTCGAGCGGTTACTTGGCCCAATGTCG	361
DB	4504	CCAGCTTTAAACTGACAGAGCGGCCCTGCTGTTTTCGAGCGGTTACTTGGCCCAATGTCG	4563
QY	362	GTGTCCTTTTCATCTTCCGAGAAAGGAAGATGTCTATTTTAAGTGACCAAGCTCAATCATG	421
DB	4564	GTGTCCTTTTCATCTTCCGAGAAAGGAAGATGTCTATTTTAAGTGACCAAGCTCAATCATG	4623
QY	422	CAAGTATGATGACGCGTGGCGATTTCTAAGGCTGATACAGTTGTTTATTCGGCATATTG	481
DB	4624	CAAGTATGATGACGCGTGGCGATTTCTAAGGCTGATACAGTTGTTTATTCGGCATATTG	4683
QY	482	ATATGAATGATCTTGAAAAACAAGCTGAATGAAAAACAAGCGTTTATCAGCGCGGTTTATCG	541
DB	4684	ATATGAATGATCTTGAAAAACAAGCTGAATGAAAAACAAGCGTTTATCAGCGCGGTTTATCG	4743
QY	542	TAAACAGACGAGTATTTCAGCATGGATGGCAAAATCGCCCTCTTGTATCAGATCATCTCAC	601
DB	4744	TAAACAGACGAGTATTTCAGCATGGATGGCAAAATCGCCCTCTTGTATCAGATCATCTCAC	4803
QY	602	TTGCGAAACGCTATCATGCTTTCGTGTCGTGTCGTGATGATGCCACGCAACAGGAGTTTGG	661
DB	4804	TTGCGAAACGCTATCATGCTTTCGTGTCGTGTCGTGATGATGCCACGCAACAGGAGTTTGG	4863
QY	662	GCGATTCCGGACAAGGAAACAGAGTGAATCTTTGGTGTGTTGTCGCCGACATTTGTTATCGGCA	721
DB	4864	GCGATTCCGGACAAGGAAACAGAGTGAATCTTTGGTGTGTTGTCGCCGACATTTGTTATCGGCA	4923

Qy	722	CCTTAAGCAAAAGCTGTGTGGCGCGAAGGAGGTTTTTTCGGCAGGATCAGCGGTCTTCATCG	781
Db	4924	CCTTAAGCAAAAGCTGTGTGGCGCGAAGGAGGTTTTTTCGGCAGGATCAGCGGTCTTCATCG	4983
Qy	782	ACTTTTTGCTGAACCATGCCAGAACATTTATCTTTCAAAACCGCTATTTCCGCGACCGAGCT	841
Db	4984	ACTTTTTGCTGAACCATGCCAGAACATTTATCTTTCAAAACCGCTATTTCCGCGACCGAGCT	5043
Qy	842	GTGCGGCTGTCTACGAGGCTTTTCAACATCATTTGAAGCCAGCAGGGAAGAAACGACAGCTTT	901
Db	5044	GTGCGGCTGTCTACGAGGCTTTTCAACATCATTTGAAGCCAGCAGGGAAGAAACGACAGCTTT	5103
Qy	902	TATTTTCTTATATCAGCATGATCAGAAACAGTCTGAAGATATGGGTATGTGGTGAAG	961
Db	5104	TATTTTCTTATATCAGCATGATCAGAAACAGTCTGAAGATATGGGTATGTGGTGAAG	5163
Qy	962	GAGATCACACACCGGATTTTCCGTAGTCAATTCGGCGATGCCATATAAAACCGTCTCTATTG	1021
Db	5164	GAGATCACACACCGGATTTTCCGTAGTCAATTCGGCGATGCCATATAAAACCGTCTCTATTG	5223
Qy	1022	CTGAAAAAAGCTGACAGGCGCAAGGGAATTTATGCTCCTGTCATTTCCGCGCGCCCAACCGTTGCGC	1081
Db	5224	CTGAAAAAAGCTGACAGGCGCAAGGGAATTTATGCTCCTGTCATTTCCGCGCGCCCAACCGTTGCGC	5283
Qy	1082	CGGGTGAAGCCGCGATTCGAAAGCTTGGGCGACAGGTCGAGATCAGGGAATGAGTTTATAA	1141
Db	5284	CGGGTGAAGCCGCGATTCGA-----ATTA	5307
Qy	1142	AATAAAAAAGACCTGAAAGGTCCTTTTTTGTATGGTTTGAACCTGCTCTTTCTTCTTA	1201
Db	5308	CAATCACTGTGACCAACAGTATGGGTGATATGATCATTTTGTGCAACAACTTCATTCAA	5367
Qy	1202	TCCTTGATACATATAGAAAATAACGTCATTTTATTTTATTTTATTTAGTTGCTCGAAAGGTCGT	1261
Db	5368	TCGGAAGAGGCTGCACATCATTTTGGGGTGTGTCACGGGAAGTCA-----	5418
Qy	1262	TGAAGTGTGGTATGATGTTTTTAAAGTATTTAAAGTATTTAAACCCCTTAAATTCGGTGCAAGA	1321
Db	5419	-----TACAGAAAGTAGGGAACCGTTATATCCACGCGCTCTG-	5456
Qy	1322	AAACCCCATCTGTTAAAGTTATAAGTGACTAAACAAATACTAAATAGATGGGGTTTC	1381
Db	5457	-----CTGCCCTTATGAAAGACAATAATAGACATCTCGGGGTGTA	5496
Qy	1382	TTTTTAATATATGTGCTCTAATAGTAGCATTTATTCAGATGAAATAATCAAGGGTTTTGTAGT	1441
Db	5497	TAAACCATTTTAAAGCGGGATATCGCGCCATCATCCAGATAGTATACAACTTTGCTGAA	5556
Qy	1442	GGACAAACAAAAGTGGAAAAGTGAGACCATGTGCTTAGGAAGACGAGTTATTTAATPAGC	1501
Db	5557	AGATATGTCGAGA-----	5570
Qy	1502	TGAATAAGAACGCTGCTCTCCAATATTTCTTATTTAGAAAGCAATCTAAATATATCTG	1561
Db	5571	-----CCAGTCTTCTCATGAAGACATTTACGCTTTTGTGCTTCAA	5610
Qy	1562	AAAAGGAATGACAATAGTGAATGGACCAATAATAATGACTAGAGAAGAAAGATGAAGA	1621
Db	5611	GGCGCGCTTGCACCATACATGTTCCAGGGAACCTTGGGGGAAGACGTCCACCATGGGAAGA	5670
Qy	1622	TTGTTTCATGAAATTAAGGAACGAATATTTGGATAAATATGGGGATGATGTTTAAAGGCTATTG	1681
Db	5671	GGTTTTAAGCCAATGGGGCGGATTAGAGAAAA-----CATGAATGCTTCATCGTAG	5723
Qy	1682	GTGTTTATGGCTCTTGTGTCGTGACATGATGGGCCCTTATTCGGATATTTGAGATGATGT	1741
Db	5724	AAGTTCGACGCGGTATTTCTGTGCCATTTGGGAGAGGACTATTTTGGTCAGTCATGTCTATAA	5783
Qy	1742	GTGTCATGTCAACAGAGGAAGCAGAGTTTCCAGCCATGAATCGACAAACCGGTGAGTGAAGG	1801
Db	5784	-----AAGCGTTGACGCTTCCATGATTTATTTGGCGCGGTCTCTCGCTTGGAAACC	5833
Qy	1802	TGGAAGTGAATTTTGTAGCGAAGAGATTTCTACTAGATTATGCATCTCAGGTGGAATCAG	1861



Db 5834 ATTAATCATACCTTTTAACTGTCAAAATATGCAGAAAGCATGGGCTCTCCAAATCGCCGA 5893  
Qy 1862 ATTGGCCGCTTACACATGGTCAATTTTCTCTATTTTGGCCGATTTATGATTCAGGTGGAT 1921  
Db 5894 ATTATCATCAATGGAATCAGTGACTCTCTCTGATGAAGATGAAAAACCAATCCTGAGATG 5953  
Qy 1922 ACTTAGAGAAAGTGTATCAAACTGCTAAATCGGTAGAGCCCAACGTTCCACAGTCGA 1981  
Db 5954 ATTGAGCGCTTATGCGGTGTCGCGATTTTAGGGTTACGCCAAGCTTGCACACGTG--- 6010  
Qy 1982 TTTGTGCCCTTATCGTAGAGAGCTGTTTGAATATGCAGGCAAAATGGCGTAAATATTCGTG 2041  
Db 6011 -----ACGAAGAAACGGTTCTACATATGTTGTAAGAACCAT- 6047  
Qy 2042 TGCAGGACCGACAAACATTTTACACATCTTGACTGTGTACAGGTAGCAATGCGAGTGC 2101  
Db 6048 -----TCAATCTATCATTTACTGATGAAATCAAGTGGGGGTATGA 6085  
Qy 2102 TGTGATTGGTCTGCATCATCGCATCTGTTATACGAGGCGCTTCGGTCTTTAACTGAAG 2161  
Db 6086 GAATGAATCAATGATGAATCGCAGACCGGTGCTGGCTGGAGCAGAAAGTGAATGACG 6145  
Qy 2162 CAGTTAAGCAATCAGATCTTCTTCAGGTTATGACCAATCTGTGCCAGTTCGTAATGCTG 2221  
Db 6146 AAGAGGCGCTTCAATATTTACATTTGCTCTGATGAAGATATTTGCTATTAATGACAGGG 6205  
Qy 2222 GTCAACTTTCGACTCTGAGAAACCTTCGGAATCGCTAGAGAAATTCGGAATGGGATTC 2281  
Db 6206 CTTTTCACATCAGAAACACTTTTACGGAAAAAAGTAAAGCTCAATATGATGATGAATG 6265  
Qy 2282 AGGAGTGACAGAACGACACGGATATATAGTGGATGTGTCAAAAGCGCATACCATTTTGA 2341  
Db 6266 CGAATCGGGCTCTGCCCGAANAACCTCGGCTATGTTTACA-----GTCTGCGAT 6317  
Qy 2342 TTCGAAAGCGCGATTGAGTCTTACCGGATGGTGAATAAGGAAACGCTGCTTGAAGCGC 2401  
Db 6318 TTCGAAAGCGCGATTGAGTCTTACCGGATGGTGAATAAGGAAACGCTGCTTGAAGCGC 6377  
Qy 2402 GAAGCGGCGCAGCATCTGAATATCGGCACATATGATTCGTGCGAAGCGCAGAGTCC 2461  
Db 6378 GAAGCGGCGCAGCATCTGAATATCGGCACATATGATTCGTGCGAAGCGCAGAGTCC 6437  
Qy 2462 GTCTAACAGAGAGTGGATCAGTCTGATGTCGGTTCAGGAAATTAAGAGACGTATGG 2521  
Db 6438 GTCTAACAGAGAGTGGATCAGTCTGATGTCGGTTCAGGAAATTAAGAGACGTATGG 6497  
Qy 2522 ACTGAAGATTTGTGCATCTCTTGGACTGTTGAAGCCAGAGCGAAGCGGCTCAAAGA 2581  
Db 6498 ACTGAAGATTTGTGCATCTCTTGGACTGTTGAAGCCAGAGCGAAGCGGCTCAAAGA 6557  
Qy 2582 TGCAGGATAGCCGCTATATCATATTTGAATACGTACAGAGAAACCAATTCAAACAT 2641  
Db 6558 TGCAGGATAGCCGCTATATCATATTTGAATACGTACAGAGAAACCAATTCAAACAT 6617  
Qy 2642 CACAACCTCACATACATACGATGACAGAGTCAATACCGTTCAAAATCGCAAAAGAAATCGG 2701  
Db 6618 CACAACCTCACATACATACGATGACAGAGTCAATACCGTTCAAAATCGCAAAAGAAATCGG 6677  
Qy 2702 GCTGTCTCCGTGTTCCAGCGCCATTTATCGGATGAAGGAGACGAAACAGGATGTCATGA 2761  
Db 6678 GCTGTCTCCGTGTTCCAGCGCCATTTATCGGATGAAGGAGACGAAACAGGATGTCATGA 6737  
Qy 2762 CATGCCAAAAGCTTGAAGGCTCTTGACCGGATCCATTCCTGTGAATTTTTTGCATGC 2821  
Db 6738 CATGCCAAAAGCTTGAAGGCTCTTGACCGGATTCATTCCTGTGAATTTTTTGCATGC 6797  
Qy 2822 AATTGATGGCAGCGCTTAGAAGCGGTCAACGAATTTAAACCCGCTGTATTCGTTAAAGT 2881  
Db 6798 AATTGATGGCAGCGCTTAGAAGCGGTCAACGAATTTAAACCCGCTGTATTCGTTAAAGT 6857  
Qy 2882 GCTGGCGCTGTTCGTTTTTATCAATCCATCAAAAGAAATTCGCAATTTCCGAGGAAGAGA 2941  
|||||

Db 6858 GCTGGCGCTGTTCGTTTTTATCAATCCATCAAAAGAAATTCGCATTTCCGAGGAAGAGA 6917  
Qy 2942 GGTCAAATCTCGCACATTTGAGCCATTTAGGCTTTTACGCGCAACCTCCATTTTGTGCG 3001  
Db 6918 GGTCAATCTCGCACATTTGAGCCATTTAGGCTTTTACGCGCAACCTCCATTTTGTGCG 6977  
Qy 3002 AGACTACTTAACTCTCCCGGCAAGAGGAGAGCGAGGATCATAAATTCGTGAGTGATTT 3061  
Db 6978 AGACTACTTAACTCTCCCGGCAAGAGGAGAGCGAGGATCATAAATTCGTGAGTGATTT 7037  
Qy 3062 AGGCTTTGAATGAAATCAGTCGAGAAATGAAGGCTAGTTTAAAGTCGAAAAGCTGAAA 3121  
Db 7038 AGGCTTTGAATGAAATCAGTCGAGAAATGAAGGCTAGTTTAAAGTCGAAAAGCTGAAA 7097  
Qy 3122 GAATCAATAAAGCAATCGGTATGATGTCGAAT 3154  
Db 7098 GAATCAATAAAGCAATCGGTATGATGTCGAT 7130

RESULT 7  
BSUB0016/c

LOCUS  
DEFINITION  
BSUB0016 199922 bp DNA linear BCT 18-APR-2005  
Bacillus subtilis complete genome (section 16 of 21): from 3013458 to 3213379.

ACCESSION  
VERSION  
Z99119 AL009126  
Z99119.2 GI:32468809

## KEYWORDS

SOURCE  
Bacillus subtilis subsp. subtilis str. 168

## ORGANISM

Bacillus subtilis subsp. subtilis str. 168

Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

## REFERENCE

1 (bases 1 to 199922)

## AUTHORS

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azevedo, V., Bertero, M.G., Bessieres, P., Bolotin, A., Borchert, S., Borriss, R., Boursiet, L., Brans, A., Braun, M., Brignell, S.C., Bron, S., Brouillet, S., Bruschi, C.V., Caldwell, B., Capuano, V., Carter, N.M., Choi, S.K., Codani, J.J., Connerton, I.F., Cummings, N.J., Daniel, R.A., Denizot, F., Devine, K.M., Dusterhoft, A., Ehrlich, S.D., Emmerger, P.T., Entian, K.D., Errington, J., Fabret, C., Ferrari, E., Foulger, D., Fritz, C., Fujita, M., Fujita, Y., Fuma, S., Galizzi, A., Galleron, N., Ghim, S.Y., Glaser, P., Goffeau, A., Golightly, E.J., Grandi, G., Guibeppe, G., Guy, B.J., Haga, K., Halech, J., Harwood, C.R., Henaut, A., Hilbert, H., Holsappel, S., Hosono, S., Hullo, M.F., Itaya, M., Jones, L., Joris, B., Karamata, D., Kasahara, Y., Klaere-Blanchard, M., Klein, C., Kobayashi, Y., Koetter, P., Koningsstein, G., Krogh, S., Kumano, M., Kurita, K., Lapidus, A., Lardinois, S., Lauber, J., Lazarevic, V., Lee, S.M., Levine, A., Liu, H., Masuda, S., Mauel, C., Medigue, C., Medina, N., Mellado, R.P., Mizuno, M., Moesti, D., Nakai, S., Noback, M., Noone, D., O'Reilly, M., Ogawa, K., Ogiwara, A., Oudega, B., Park, S.H., Parro, V., Pohl, T.M., Portetelle, D., Porwollik, S., Prescott, A.M., Prescann, E., Pujic, P., Purnelle, B., Rapoport, G., Rey, M., Reynolds, S., Rieger, M., Rivolta, C., Rocha, E., Roche, B., Rose, M., Sadaie, Y., Sato, T., Scanlan, E., Schleich, S., Schroeter, R., Scoffone, F., Sekiguchi, J., Sekowska, A., Seror, S.J., Serrero, P., Shin, B.S., Soldo, B., Sorokin, A., Tacconi, E., Takagi, T., Takahashi, H., Takemaru, K., Takeuchi, M., Takakoshi, A., Tanaka, T., Terpetra, P., Tognoni, A., Tosato, V., Uchiyama, S., Vandenbol, M., Vannier, F., Vaasarotti, A., Viari, A., Wambutt, R., Wedler, E., Wedler, H., Weitzengraber, T., Winters, P., Wipat, A., Yamamoto, H., Yamane, K., Yasumoto, K., Yata, K., Yoshida, K., Yoshikawa, H.F., Zumstein, E., Yoshikawa, H. and Danchin, A.

The complete genome sequence of the gram-positive bacterium

## TITLE

Bacillus subtilis

Nature 390 (6657), 249-256 (1997)

9384377

2 (bases 1 to 199922)

## REFERENCE

Kunst, F., Ogasawara, N., Yoshikawa, H. and Danchin, A.

## AUTHORS

Submitted (27-JUN-2003) I. Moszer, A. Danchin, Institut Pasteur,

Genetique des Genomes Bacteriens, 28 rue du Docteur Roux, 75224

Paris Cedex 15, FRANCE. E-mail: moszer@pasteur.fr,

adanchin@pasteur.fr Phone: +33 (0)1 45 68 84 41, Fax: +33 (0)1 45

68 89 48

## COMMENT

On Jul 7, 2003 this sequence version replaced gi.2635411.  
This entry contains data from release R16.1 of the Subtilist  
database. Further data on gene annotation and detailed information  
about changes from previous releases can be found at  
<http://genolist.pasteur.fr/Subtilist/>.

## FEATURES

## source

1. 199922

/organism="Bacillus subtilis subsp. subtilis str. 168"

/mol\_type="genomic DNA"

/strain="168"

/db\_xref="taxon:224308"

/complement(96..608)

/gene="moaB"

/locus\_tag="BSU29460"

/complement(96..608)

/gene="moaB"

/locus\_tag="BSU29460"

/codon\_start=1

/transl\_table=11

/product="molybdopterin precursor biosynthesis protein B"

/protein\_id="CAB14924.1"

/db\_xref="GI:2635430"

/db\_xref="GOA:O34457"

/db\_xref="UniProt:IPR001453"

/db\_xref="InterPro:IPR008284"

/db\_xref="Subtilist:BG12618"

/db\_xref="UniProt/Swiss-Prot:O34457"

/translation="MSVQEHKQEPADIVRCVITVSDTRTDETDKSKLMISPLEEAG

HHIAAYIVKDEKALRSVLGMSQDQTDVALLNGTGADRDVTEATPTFSKEL

FGGEIFPMLSYEDIGSAAIMSRATAGVIQHTAVFTPGSGGAVKLANMKLIPELA

HVVRIRKDK"

/complement(693..1880)

/gene="ackA"

/locus\_tag="BSU29470"

/complement(693..1880)

/gene="ackA"

/locus\_tag="BSU29470"

/EC number="2.7.2.1"

/function="conversion of acetyl-CoA to acetate"

/codon\_start=1

/transl\_table=11

/product="acetate kinase"

/protein\_id="CAB14925.1"

/db\_xref="GI:2635431"

/db\_xref="GOA:P37877"

/db\_xref="InterPro:IPR000890"

/db\_xref="InterPro:IPR004372"

/db\_xref="Subtilist:BG10813"

/db\_xref="UniProt/Swiss-Prot:P37877"

/translation="MSKIIAIINAGSSSLKQLFEMPSETVLTGKLVIRIGIADSVFTI

SVNGKTEVTDIPDAVAVKMLNKLTEFGIILKDLNEIDGIGHRVHVGKEFSDSVL

LTDETIKEIDISLAPLHPNPANTVGIKAFKVLPNPAVAFDTAFHPTMPQSVLY

SLPYEYKRGIRGIFGHGSHKIVTERRAELGRPLKDLRLISCHLNGASTAAVEG

SKSDTSWGTPLAGVAMGTRGNIDPALIPIYMEKGTQDADEVLTNKKSGLLGIS

GFSSDLDIYEATKEGNEARATALEVFASRIHKYIGSYAARMSGVDAIIFTAGIGNS

VEVREVRILRGLEFGVWDPALNVNVRGEAFISYPHSPVKVMIPTDEVMIAQDVR

LAK"

/complement(2228..3217)

/gene="ytxK"

/locus\_tag="BSU29480"

/complement(2228..3217)

/gene="ytxK"

/locus\_tag="BSU29480"

/function="unknown"

/note="alternate gene name: ythI

similar to unknown proteins"

/codon\_start=1

/transl\_table=11

/protein\_id="CAB14926.1"

/db\_xref="GI:2635432"

/db\_xref="GOA:P37876"

/db\_xref="InterPro:IPR000051"

/db\_xref="InterPro:IPR002296"

## gene

/db\_xref="Subtilist:BG10812"

/db\_xref="UniProt/Swiss-Prot:P37876"

/translation="MOKDHVAVYELLNEAAIMIKNELQISYIEALAEAGEMYFLEKT

DOLKLPADOKTKQALLKAEFGTYEHWRKAFQALVLMGKMDISHPNRQMTPTDI

GLFISLVNKKPMADKXELTILDPALGTGNLLFTVLNQLSEKTSANSFGIETDVLKIA

YAQNLLKLELPHQDSLEPLFIDPDVTVICDLPVGYIYFNDEGAFAFELKADGHSF

AHLHFTQSVKHTPGGYLPFMPINHLFESSQSGKLQFFKDKVHINALLQLPKSIPIK

DEAHAKSILVLQFGENTKAPQGIILLANLPSFSNQKAMLDMAQDFDEWFKKEK"

/complement(3278..3781)

/gene="tpx"

/locus\_tag="BSU29490"

/complement(3278..3781)

/gene="tpx"

/locus\_tag="BSU29490"

/note="alternate gene name: ytgI"

/codon\_start=1

/transl\_table=11

/product="thiol peroxidase"

/protein\_id="CAB14927.1"

/db\_xref="GI:2635433"

/db\_xref="GOA:P80864"

/db\_xref="Subtilist:BG13855"

/db\_xref="UniProt/Swiss-Prot:P80864"

/translation="MAEITPKGGPVLVGGVKGQADPFTVLTNSLEKSLADMKG

KVTIISVPSIDTGVCDQTRFNEEAAGLDVNVYVTSADLPFAQARWCGANGIDKV

ETLSDHRDMSFGSAFGVYIKELRLARSFVLDENGKVYAYVSEATNHPNVEKPIE

AKALVK"

/complement(3891..4346)

/gene="ytfJ"

/locus\_tag="BSU29500"

/complement(3891..4346)

/gene="ytfJ"

/locus\_tag="unknown"

/function="unknown"

/note="similar to unknown proteins"

/codon\_start=1

/transl\_table=11

/protein\_id="CAB14928.1"

/db\_xref="GI:2635434"

/db\_xref="Subtilist:BG13849"

/db\_xref="UniProt/TrEMBL:O34806"

/translation="MADHPIQGLMKMTAMENLKEMIDVNTIIGDPVETPDGSLVTSK

VGFGAAGSGFCGPAEKSEDETRQKLPFGGGSGGVSIPTPIAFLIVGSTGIRM

LHLDENTHLEKILDAAPQTLERIQOMFKKNNKNSQSQGNQNNNNY"

/complement(4360..5040)

/gene="ytfI"

/locus\_tag="BSU29510"

/complement(4360..5040)

/gene="ytfI"

/locus\_tag="unknown"

/function="unknown"

/note="similar to unknown proteins"

/codon\_start=1

/transl\_table=11

/protein\_id="CAB14929.1"

/db\_xref="GI:2635435"

/db\_xref="Subtilist:BG13848"

/db\_xref="UniProt/TrEMBL:O34917"

/translation="WYVLTAILLIGVILLRMKIHVALEYLHADNDKLTAL

FGILIRKTIIPAKVKNKDGTVDIKORTESMKKSGKKITQNDVQGSINKLEMLH

QVTRLRKISASFLAHLHTKLEWTVIGIQDAAVTGLTVGAVNGKGGIAAMLYDHL

FVNKPVYEVIPSPQVPVSKTHFCIIFFRFGHAMLAFAKFKVYKRGRELSVLKQNSARM

TSKNDSSV"

/complement(5115..5609)

/gene="yteJ"

/locus\_tag="BSU29520"

/complement(5115..5609)

/gene="yteJ"

/locus\_tag="BSU29520"

/function="unknown"

/note="similar to unknown proteins"

/codon\_start=1

/transl\_table=11

/protein\_id="CAB14930.1"  
/db\_xref="GI:2635436"

Query Match 51.1%; Score 1612.6; DB 1; Length 199922;  
Best Local Similarity 73.0%; Pred. No. 0;  
Matches 2303; Conservative 0; Mismatches 624; Indels 226; Gaps 9;

QY 2 GATCCACGAGGTACGAGCCTTGAAGATTGATTCCTGGTTAAACGAGCGGTTAGACAGAA 61  
DB 78965 GATCCACGAGGTACGAGCCTTGAAGATTGATTCCTGGTTAAACGAGCGGTTAGACAGAA 78966  
QY 62 TGAAGAAGACGGCGGTACATCGTAACCTGCGGTCAATGGATGAGCGCGGTTCCAGAGA 121  
DB 78965 TGAAGAAGACGGCGGTACATCGTAACCTGCGGTCAATGGATGAGCGCGGTTCCAGAGA 78966  
QY 122 GGAATATTGATGGCGGAAATCAACGGTCTGGTCTCAAAACAATTTATGGGCTCGCAA 181  
DB 78845 GGAATATTGATGGCGGAAATCAACGGTCTGGTCTCAAAACAATTTATGGGCTCGCAA 78786  
QY 182 CGGATAGACGTTTATGATCGATCGAGCCCAACAGCAATTCGAGCAATTTGGGACAGGAAGCA 241  
DB 78785 CGGATAGACGTTTATGATCGATCGAGCCCAACAGCAATTCGAGCAATTTGGGACAGGAAGCA 78726  
QY 242 CGCGTTTCAAGTAAACGAGCGCAATTCGGTCTGGCATGAAAGCTAGAAAAGAAAGATTG 301  
DB 78725 CGCGTTTCAAGTAAACGAGCGCAATTCGGTCTGGCATGAAAGCTAGAAAAGAAAGATTG 78666  
QY 302 CCAGCTTTAAACTGACAGAGCGGCGCTGCTGTTTTCGAGCGGTTTTCGAGCGCAATGTCG 361  
DB 78665 CCAGCTTTAAACTGACAGAGCGGCGCTGCTGTTTTCGAGCGGTTTTCGAGCGCAATGTCG 78606  
QY 362 GTGCTCTTTTCACTTTCGAGAAAGGAAGATGATCTTTTAAAGTACACAGCTCAATCATG 421  
DB 78605 GTGCTCTTTTCACTTTCGAGAAAGGAAGATGATCTTTTAAAGTACACAGCTCAATCATG 78546  
QY 422 CAAGTATGATCGAGCGCTCGGACCTTTCTTAAGGCTGATACAGTGTGTTTATCGGCATATTG 481  
DB 78545 CAAGTATGATCGAGCGCTCGGACCTTTCTTAAGGCTGATACAGTGTGTTTATCGGCATATTG 78486  
QY 482 ATATGAATGATCTTGAACAAAGCTGAAATGAAACACAGCGTTATCAGCGCGGTTTATCG 541  
DB 78485 ATATGAATGATCTTGAACAAAGCTGAAATGAAACACAGCGTTATCAGCGCGGTTTATCG 78426  
QY 542 TAAACAGCGGAGTATTCAGCATGATGATGCAATCGCCCTCTTGATCAGATCATCTCAC 601  
DB 78425 TAAACAGCGGAGTATTCAGCATGATGATGCAATCGCCCTCTTGATCAGATCATCTCAC 78366  
QY 602 TTGCGAAACGCTATCATGCTTTCGTTGTCGTTGATGATGCCACGCAACAGGAGTTTGG 661  
DB 78365 TTGCGAAACGCTATCATGCTTTCGTTGTCGTTGATGATGCCACGCAACAGGAGTTTGG 78306  
QY 662 CGGATTCGGGACAGGAAACGAGTGAATCTTTGTTGTTTGGTCCGACATTTGTTATCGGCA 721  
DB 78305 CGGATTCGGGACAGGAAACGAGTGAATCTTTGTTGTTTGGTCCGACATTTGTTATCGGCA 78246  
QY 722 CCTTAAGCAAGCTGTTGCGCGGAGGAGGTTTTCGCGAGGATTCAGCGGCTTCTATCG 781  
DB 78245 CCTTAAGCAAGCTGTTGCGCGGAGGAGGTTTTCGCGAGGATTCAGCGGCTTCTATCG 78186  
QY 782 ACTTTTTCGCTGAACCATGCCAGAAATTTATCTTTTCAACCGCTATTTCCGCCAGCAGCT 841  
DB 78185 ACTTTTTCGCTGAACCATGCCAGAAATTTATCTTTTCAACCGCTATTTCCGCCAGCAGCT 78126  
QY 842 GTGCGGCTGCTCACGAGGCTTTCAACATCATTTGAAGCCAGCAGGAAACGACAGCTTT 901  
DB 78125 GTGCGGCTGCTCACGAGGCTTTCAACATCATTTGAAGCCAGCAGGAAACGACAGCTTT 78066  
QY 902 TATTTTCTTATACAGCATGATCAGAACAGCTGCAAGAAATATGTTGTTATGTTGTTGAAAG 961  
DB 78065 TATTTTCTTATACAGCATGATCAGAACAGCTGCAAGAAATATGTTGTTATGTTGTTGAAAG 78006  
QY 962 GAGATCACACCGGATTTCTGTTAGTCAATGGCGATGCCCATAAACGGTCTCTATTG 1021

DB 78005 GAGATCACACCGGATTTATCTCTGTAGTCAATGGCGATGCCCATAAACCGTCTCTATTG 77946  
QY 1022 CTGAAAACCTGCAGGCGCAAGGAAATTTATGCTCTGCAATTCGCGCGCCACCGCTTCGCG 1081  
DB 77945 CTGAAAACCTGCAGGCGCAAGGAAATTTATGCTCTGCAATTCGCGCGCCACCGCTTCGCG 77886  
QY 1082 CGGCTGAAAGCCGATTCGAAAGCTTGGGCGACAGGTCGAGATCAGGGAATGAGTTTATAA 1141  
DB 77885 CGGCTGAAAGCCGATTCGA-----ATTA 77862  
QY 1142 AATAAAAAAGCACCTGAAAAGGTGCTTTTTTTTGTGATGTTTGAATCTGTTCTTTCTTA 1201  
DB 77861 CAATCACGTCTGACCAAGTATGAGTATGATCATTTGCTGCAACATTTTCATTTCAA 77802  
QY 1202 TCTTGATACATATAGAAATAACGTCATTTTATTTTATTTTATTTAGTCTGCTGAAAGGTGCGT 1261  
DB 77801 TCGGAAAGGAGCTGCACATCATTTGAGGGGTTTTTTTGTGACGGGAACTGA----- 77751  
QY 1262 TGAAGTGTGTGATGATGTGTTTAAAGTATTTGAAAACCCCTTAAAAATTTGGTTGCACAGA 1321  
DB 77750 -----TACAGAAGTAGGAAAACGGTTATATCCAGCGGTCTG- 77713  
QY 1322 AAAACCCCATCTGTTAAAGTTATAAGTGAATAAACAATACTAAATAGATGCGGGTTTC 1381  
DB 77712 -----CTGCTTTATTTGAAAAGACAATAATAGACATGTGCGGGGTGTA 77673  
QY 1382 TTTTAAATTTATGTCCTTAATAGTAGCATTTATTCAGATGMAAAATCAAGGTTTTAGT 1441  
DB 77672 TAAACCATTTTAAAGCGGGATATCGCGCCATCATCCAGATAGTATCAAGTTTGTGAA 77613  
QY 1442 GGACAGACAAAAAGTGGAAAAAGTGAGACCATGTGCTTTAGGAAGACGAGTTATTAATAGC 1501  
DB 77612 AGATATGTCGAGA----- 77599  
QY 1502 TGAATAAGAACGGTGTCTCCAAATATCTTATTTAGAAAGCAAACTTAAAAATTTATCTG 1561  
DB 77598 -----CCAGTCTTCTCATGAAGACATTTACGCTTTTTCCTTCAA 77559  
QY 1562 AAAAGGGAATGAGAAATAGTGAATGGACCAATAATATGACTAGAGAAGAAGATGAAGA 1621  
DB 77558 GGCGCGCTTGCACCATACATGTTGCGGGGAACTTTGAGGAAAGAGCTCTCACCATTGGAAGA 77499  
QY 1622 TTGTTTCAATGAAATTAAGGAAACGAATATTGGATAAATATGGGATGATGTTTAAAGCTATTG 1681  
DB 77498 GGTTTTAAAGCCATTTGGGGCGGATTAGAGAAJAA-----CATGAATGCTTCATCGTAG 77446  
QY 1682 GTGTTTATGGCTCTCTTGGTCTGTCAGACTGATGAGGCGCTTATTCGGATATTGAGATGATG 1741  
DB 77445 AAGGTGCGGCGGTATTTCTGTGCCATTGGGAGAGGACTATTTGGTCAGTCATGTCATAA 77386  
QY 1742 GTGTCATGTCACAGAGGAAGCAGAGTTTCAGCCATGAATGGACAACCGGTGAGTGGGAAGG 1801  
DB 77385 -----AAGGTTGCGAGCTTCCCATGATTTATGTCGCGCTCTCGCTTTGGAACC 77336  
QY 1802 TGAAGTGAATTTTGTAGGAAAGAGATTTCTATAGATTTATGCAATCTCAGGTGGAATCAG 1861  
DB 77335 ATTAATCATACCTTTTAACTGTCAAAATATGCAAGAGCATGGGCTTCCAAATCGCGGA 77276  
QY 1862 ATTGGCGCTTACATAGGTCAATTTTCTCTATTTTTCGCGATTTTATGATTCAGGTGAT 1921  
DB 77275 ATTATCATCAATGGAATCAGTGAATCTCTGATGAAGATGAAAACCAATCTTGAGATG 77216  
QY 1922 ACTTAGAGAAGTGTATCAAACTGCTAAATCGGTGAGAGCCCAACGTTCCACGATGCGA 1981  
DB 77215 ATTGAGCGCTTATGCGGTGTGCCGATTTTGGGGTTTACGCCAAGCTTGCACAGCTG--- 77159  
QY 1982 TTTGTGCCCTTATCGTAGAAGAGCTGTTTGAATATGCGGCAATATGCGGTAATATTCTGT 2041  
DB 77158 -----ACGAAGAAACGGTCTCTACATATGTTAAAGACCAT- 77122  
QY 2042 TGCAGGACCGCAACAAATTTCTACCATCTTGTACTGTACAGGTAGCAATGCGAGGTGCCA 2101  
DB 77121 -----TCAATCTATCTACTGATGAATCAAGTGGGGGTATGA 77084

Qy 2102 TGTGATGGTCTGCATCATCGCATCTGTTATACGACGAGCGCTTCGGTCTTAACTGAAG 2161  
| | | | |  
Db 77083 GAATGAATCAATGATGGAATCTGCAGACCGGGTCTGGCTGGAGCAGAAGTGAAGTCAAG 77024  
Qy 2162 CAGTTAAGCAATCAGATCTCCCTTCAGTTATGACCAATCTGTGCGAGTTCGTAATGCTG 2221  
| | | | |  
Db 77023 AAGAGGCGCTTCAATATATACATTTGTCTGTAAGATATTTTGTCTATTATGATGACGCGG 76964  
| | | | |  
Qy 2222 GTCAACTTTCGACTCTGAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTC 2281  
| | | | |  
Db 76963 CTTTTCACATCAGAAACACATTTTACGGAATAAAGTAAGCTCAATATGATGATGATG 76904  
Qy 2282 AGGAGTGGACAGAACGACGCGATATATAGTGGATGTGTCAAAACGACATACCATTTTCAA 2341  
| | | | |  
Db 76903 CGAAATCGGGCTCTGCCCGGAAACTCGCGCTATTGTTTACA-----GTCTCGCAT 76852  
| | | | |  
Qy 2342 TTGCAAGCGCGGATGAGTCTTACCGGATGGTGAATAAGAAACGCTGCTTGAAGCGC 2401  
| | | | |  
Db 76851 TTGCAAGCGCGGATGAGTCTTACCGGATGGTGAATAAGAAACGCTGCTTGAAGCGC 76792  
| | | | |  
Qy 2402 GAAGCGGCGCAGCATCTGAATATCGGCACATATTTGTATCGTGCAGACGCGCAGAGTCC 2461  
| | | | |  
Db 76791 GAAGCGGCGCAGCATCTGAATATCGGCACATATTTGTATCGTGCAGACGCGCAGAGTCC 76732  
Qy 2462 GTCTAACAGAGAGTGGATCAGTGTGATGATCGGTTTCAGGAATTTAAAGAGACGTATGG 2521  
| | | | |  
Db 76731 GTCTAACAGAGAGTGGATCAGTGTGATGATCGGTTTCAGGAATTTAAAGAGACGTATGG 76672  
| | | | |  
Qy 2522 ACTGAAGATTTGTGATGCTTGTGACTGTGTAAGCGCAGAGCAGCGAAGCGGCTCAAAGA 2581  
| | | | |  
Db 76671 ACTGAAGATTTGTGATGCTTGTGACTGTGTAAGCGCAGAGCAGCGAAGCGGCTCAAAGA 76612  
| | | | |  
Qy 2582 TGCAGGAGTACGCGCTATATATATATTTTGAATACGTCACAGAGAAACCATTCAAACAT 2641  
| | | | |  
Db 76611 TGCAGGAGTACGCGCTATATATATTTTGAATACGTCACAGAGAAACCATTCAAACAT 76552  
Qy 2642 CACAACCTCACATACATACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2701  
| | | | |  
Db 76551 CACAACCTCACATACATGATGATGATGATGATGATGATGATGATGATGATGATGATG 76492  
Qy 2702 GCTGTCTCCGTGTTTCAGCGCGCATTTTCGGATGAAGAGACGAAACAGGATGATGATGA 2761  
| | | | |  
Db 76491 GCTGTCTCCGTGTTTCAGCGCGCATTTTCGGATGAAGAGACGAAACAGGATGATGATGA 76432  
Qy 2762 CATCGCCAAAGCTTGAAGGCTCTTGAAGCGGATTCATTCCTGTAATTTTTCGATGC 2821  
| | | | |  
Db 76431 CATCGCCAAAGCTTGAAGGCTCTTGAAGCGGATTCATTCCTGTAATTTTTCGATGC 76372  
Qy 2822 AATTGATGGCAGCGCTTGAAGCGGCTCAACGAATTTAAACCCGCTGTATTTTAAAGT 2881  
| | | | |  
Db 76371 AATTGATGGCAGCGCTTGAAGCGGCTCAACGAATTTAAACCCGCTGTATTTTAAAGT 76312  
Qy 2882 GCTGGCGTGTTCGTTTTATCAATTCATCAAAAGAAATTCGCATTTTCGGAGAGAGAGA 2941  
| | | | |  
Db 76311 GCTGGCGTGTTCGTTTTATCAATTCATCAAAAGAAATTCGCATTTTCGGAGAGAGAGA 76252  
Qy 2942 GGTCAATCTCGCACATTCGAGCCATTTAGGCTTTTACGCGCAACCTCCATTTTGTGCG 3001  
| | | | |  
Db 76251 GGTCAATCTCGCACATTCGAGCCATTTAGGCTTTTACGCGCAACCTCCATTTTGTGCG 76192  
Qy 3002 AGACTACTTAACTGCGGCGCAAGAGAGAGAGAGAGATCATAAATTCGTGAGTGAATTT 3061  
| | | | |  
Db 76191 AGACTACTTAACTGCGGCGCAAGAGAGAGAGAGATCATAAATTCGTGAGTGAATTT 76132  
Qy 3062 AGGCTTTGAAGTTGAATCAGTCGAGAAATTAAGAGGCTAGTTTAAAGTGGAAAGCTGAAA 3121  
| | | | |  
Db 76131 AGGCTTTGAAGTTGAATCAGTCGAGAAATTAAGAGGCTAGTTTAAAGTGGAAAGCTGAAA 76072  
Qy 3122 GAATCAATAAAGCAATCGGTATCATGTCGAT 3154  
| | | | |  
Db 76071 GAATCAATAAAGCAATCGGTATCATGTCGAT 76039

RESULT 8  
AF008220  
LOCUS 22060 bp DNA linear BCT 04-FEB-1998  
DEFINITION Bacillus subtilis rrnB-dnaB genomic region.  
ACCESSION AF008220  
VERSION AF008220.1 GI:2293135  
KEYWORDS  
SOURCE  
ORGANISM Bacillus subtilis  
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
REFERENCE 1 (bases 1 to 7430)  
AUTHORS Green,C.J., Stewart,G.C., Hollis,M.A., Vold,B.S. and Bott,K.F.  
TITLE Nucleotide sequence of the Bacillus subtilis ribosomal RNA operon, rrnB  
Gene 37 (1-3), 261-266 (1985)  
PUBMED 2414156  
REFERENCE 2 (bases 153210 to 153762)  
AUTHORS Connors,M.J., Mason,J.M. and Setlow,P.  
TITLE Cloning and nucleotide sequencing of genes for three small, acid-soluble proteins from Bacillus subtilis spores  
J. Bacteriol. 166 (2), 417-425 (1986)  
PUBMED 3009398  
REFERENCE 3 (bases 213161 to 218473)  
AUTHORS Ogasawara,N., Moriya,S., Mazza,P.G. and Yoshikawa,H.  
TITLE Nucleotide sequence and organization of dnaB gene and neighbouring genes on the Bacillus subtilis chromosome  
Nucleic Acids Res. 14 (24), 9989-9999 (1986)  
PUBMED 3027671  
REFERENCE 4 (bases 200404 to 201481)  
AUTHORS Saki,T., Yoshikawa,H., Takahashi,H. and Saito,H.  
TITLE Cloning and nucleotide sequence of phoP, the regulatory gene for alkaline phosphatase and phosphodiesterase in Bacillus subtilis  
J. Bacteriol. 169 (7), 2913-2916 (1987)  
PUBMED 3036763  
REFERENCE 5 (bases 201234 to 203212)  
AUTHORS Saki,T., Yoshikawa,H., Takahashi,H. and Saito,H.  
TITLE Nucleotide sequence of the Bacillus subtilis phoR gene  
J. Bacteriol. 170 (12), 5935-5938 (1988)  
PUBMED 3142862  
REFERENCE 6 (bases 142232 to 144147)  
AUTHORS Grundy,F.J. and Henkin,T.M.  
TITLE Cloning and analysis of the Bacillus subtilis rpsD gene, encoding ribosomal protein S4  
J. Bacteriol. 172 (11), 6372-6379 (1990)  
PUBMED 1699930  
REFERENCE 7 (bases 133624 to 134990)  
AUTHORS Henkin,T.M., Grundy,F.J., Nicholson,W.L. and Chambliss,G.H.  
TITLE Catabolite repression of alpha-amylase gene expression in Bacillus subtilis involves a trans-acting gene product homologous to the Escherichia coli lacI and galP repressors  
Mol. Microbiol. 5 (3), 575-584 (1991)  
PUBMED 1304524  
REFERENCE 8 (bases 140810 to 142610)  
AUTHORS Henkin,T.M., Glass,B.L. and Grundy,F.J.  
TITLE Analysis of the Bacillus subtilis tyrS gene: conservation of a regulatory sequence in multiple trRNA synthetase genes  
J. Bacteriol. 174 (4), 1299-1306 (1992)  
PUBMED 1735721  
REFERENCE 9 (bases 217570 to 220060)  
AUTHORS Putzer,H., Gendron,N. and Grunberg-Manago,M.  
TITLE Co-ordinate expression of the two threonyl-tRNA synthetase genes in Bacillus subtilis: control by transcriptional antitermination involving a conserved regulatory sequence  
EMBO J. 11 (8), 3117-3127 (1992)  
PUBMED 1379177  
REFERENCE 10 (bases 134990 to 141290)  
AUTHORS Grundy,F.J., Waters,D.A., Takova,T.Y. and Henkin,T.M.  
TITLE Identification of genes involved in utilization of acetate and acetoin in Bacillus subtilis  
Mol. Microbiol. 10 (2), 259-271 (1993)  
PUBMED 7934817  
REFERENCE 11 (bases 162129 to 164080)

AUTHORS TITLE JOURNAL PUBMED REFERENCE AUTHORS TITLE	Grundy, F.J., Waters, D.A., Allen, S.H. and Henkin, T.M. Regulation of the Bacillus subtilis acetate kinase gene by CcpA J. Bacteriol. 175 (22), 7348-7355 (1993) 8226682 12 (bases 7009 to 15526) Kiel, J.A., Boels, J.M., Beldman, G. and Venema, G. Glycogen in Bacillus subtilis: molecular characterization of an operon encoding enzymes involved in glycogen biosynthesis and degradation Mol. Microbiol. 11 (1), 203-218 (1994) 8145641 13 (bases 196487 to 200620) Jin, S. and Sonenshein, A.L. Identification of two distinct Bacillus subtilis citrate synthase genes J. Bacteriol. 176 (15), 4669-4679 (1994) 8045898 14 (bases 16985 to 19588) Abe, A., Koide, H., Kohn, T. and Watabe, K. A Bacillus subtilis spore coat polypeptide gene, cots Microbiology 141 (Pt 6), 1433-1442 (1995) 7545510 15 (bases 131934 to 133970) Bolotin, A., Khazak, V., Stoyanova, N., Ratmanova, K., Yomantas, Y. and Kozlov, Y. Identical amino acid sequence of the aroA(G) gene products of Bacillus subtilis 168 and B. subtilis Marburg strain Microbiology 141 (Pt 9), 2219-2222 (1995) 7496534 16 (bases 25258 to 31212) Rowland, B., Hill, K., Miller, P., Driscoll, J. and Taber, H. Structural organization of a Bacillus subtilis operon encoding menaquinone biosynthetic enzymes Gene 167 (1-2), 105-109 (1995) 8566759 17 (bases 196487 to 200620) Jin, S., De Jesus-Berrios, M. and Sonenshein, A.L. A Bacillus subtilis malate dehydrogenase gene J. Bacteriol. 178 (2), 560-563 (1996) 8550482 18 (bases 129888 to 132207) Varon, D., Brody, M.S. and Price, C.W. Bacillus subtilis operon under the dual control of the general stress transcription factor sigma B and the sporulation transcription factor sigma H Mol. Microbiol. 20 (2), 339-350 (1996) 873232 19 (bases 81540 to 91690) Bower, S., Perkins, J.B., Yocum, R.R., Howitt, C.L., Rahaim, P. and Pero, J. Cloning, sequencing, and characterization of the Bacillus subtilis biotin biosynthetic operon J. Bacteriol. 178 (14), 4122-4130 (1996) 8763940 20 (bases 49093 to 51682) Yocum, R.R., Perkins, J.B., Howitt, C.L. and Pero, J. Cloning and characterization of the metE gene encoding S-adenosylmethionine synthetase from Bacillus subtilis J. Bacteriol. 178 (15), 4604-4610 (1996) 8755891 21 (bases 100760 to 102298) Kappes, R.M., Kempf, B. and Bremer, E. Three transport systems for the osmoprotectant glycine betaine operate in Bacillus subtilis: characterization of OpuB J. Bacteriol. 178 (17), 5071-5079 (1996) 8752321 22 (bases 1 to 220060) Lapidus, A., Galleron, N., Sorokin, A. and Ehrlich, S.D. Sequencing and functional annotation of the Bacillus subtilis genes in the 200 kb rnb-dnaB region Microbiology 143 (Pt 11), 3431-3441 (1997) 9387221 23 (bases 1 to 220060)	AUTHORS TITLE JOURNAL PUBMED REFERENCE AUTHORS TITLE	
AUTHORS TITLE JOURNAL PUBMED REFERENCE AUTHORS TITLE	Lapidus, A., Galleron, N., Sorokin, A. and Ehrlich, D. Direct Submission Submitted (13-JUN-1997) Laboratoire de Genetique Microbienne, INRA Domaine de Vilvert, Jouy-en-Josas cedex 78352, France Location/Qualifiers 1. 220060 /organism="Bacillus subtilis" /mol_type="genomic DNA" /db_xref="taxon:1423" 485. .2034 /product="16S ribosomal RNA" 2203. .5129 /product="23S ribosomal RNA" 5185. .5300 /product="5S ribosomal RNA" 5322. .5397 /product="tRNA-Val" /note="codon recognized: GUA" 5430. .5504 /product="tRNA-Thr" /note="codon recognized: ACA" 5543. .5618 /product="tRNA-Lys" /note="codon recognized: AAA" 5629. .5715 /product="tRNA-Leu" /note="codon recognized: CUG" 5721. .5795 /product="tRNA-Gly" /note="codon recognized: GGC" 5810. .5895 /product="tRNA-Leu" /note="codon recognized: UUA" 5905. .5981 /product="tRNA-Arg" /note="codon recognized: CGU" 5997. .6073 /product="tRNA-Pro" /note="codon recognized: CCA" 6079. .6151 /product="tRNA-Ala" /note="codon recognized: GCA" 6172. .6248 /product="tRNA-Met" /note="codon recognized: AUG" 6251. .6327 /product="tRNA-Ile" /note="codon recognized: AUG" 6334. .6425 /product="tRNA-Ser" /note="codon recognized: UCA" 6443. .6519 /product="tRNA-Met" /note="codon recognized: AUG" 6531. .6607 /product="tRNA-Asp" /note="codon recognized: GAC"	AUTHORS TITLE JOURNAL PUBMED REFERENCE AUTHORS TITLE	
Query Match Best Local Similarity Matches 2303; Conservative	51.1%; Score 1612.6; DB 1; Length 220060; 73.0%; Pred. No. 0; 0; Mismatches 624; Indels 226; Gaps 9;	Query Match Best Local Similarity Matches 2303; Conservative	51.1%; Score 1612.6; DB 1; Length 220060; 73.0%; Pred. No. 0; 0; Mismatches 624; Indels 226; Gaps 9;
Qy	2 GATCCAGGTTACGAGCCTTGAAGATTGATTCCTGGTTAAACGAGCGGTACACAGAA 61	Qy	62 TGAAGAAGCGCGGTACATCGTAACTCGGTCAATGGATGGAGCGCGGTTCACAGAA 121
Db	85743 GATCCAGGTTACGAGCCTTGAAGATTGATTCCTGGTTAAACGAGCGGTACACAGAA 85803	Db	85803 TGAAGAAGCGCGGTACATCGTAACTCGGTCAATGGATGGAGCGCGGTTCACAGAA 85863
Qy	122 GGAATATTGATGGCGAAAATCAAAACGGTCTGGTCTCTCAACAATATTATTAGGCTCGCAA 181	Qy	122 GGAATATTGATGGCGAAAATCAAAACGGTCTGGTCTCTCAACAATATTATTAGGCTCGCAA 181
Db	85863 GGAATATTGATGGCGAAAATCAAAACGGTCTGGTCTCTCAACAATATTATTAGGCTCGCAA 85922	Db	85863 GGAATATTGATGGCGAAAATCAAAACGGTCTGGTCTCTCAACAATATTATTAGGCTCGCAA 85922



Db 87857 TTCCGAAGCGCGGATTGAGTCTTACCGCATGTTGAATAAGGAACCGTCTGTTGAAGCGC 87916  
Qy 2402 GAAGCGGCGCAGCATCTGAATATCGGCACATATTTGATCGTGCAGACGCGCAGAGTCC 2461  
Db 87917 GAAGCGGCGCAGCATCTGAATATCGGCACATATTTGATCGTGCAGACGCGCAGAGTCC 87976  
Qy 2462 GTCTAAACAGAGAAGTGGATCAGGTGATGTCGGTTCAGGAAATTTAAAGAGACGTATGG 2521  
Db 87977 GTCTAACAGAGAAGTGGATCAGGTGATGTCGGTTCAGGAAATTTAAAGAGACGTATGG 88036  
Qy 2522 ACTGAAGATTGTGCATGCTTGGACGTGTTGAAGCCAGACGAGCGCGCTCAAAGA 2581  
Db 88037 ACTGAAGATTGTGCATGCTTGGACGTGTTGAAGCCAGACGAGCGCGCTCAAAGA 88096  
Qy 2582 TGCAGGAGTAGACCGCTATATCATATTTGAATACGTACAGAGAAACCATTTCAAACAT 2641  
Db 88097 TGCAGGAGTAGACCGCTATATCATATTTGAATACGTACAGAGAAACCATTTCAAACAT 88156  
Qy 2642 CACAACCTCACATACATACGATGACAGATCAATACGGTTGAAATCGCAAAAGAAATCGGG 2701  
Db 88157 CACAACCTCACATACATACGATGACAGATCAATACGGTTGAAATCGCAAAAGAAATCGGG 88216  
Qy 2702 GCTGTCTCCGTGTTCAGCGCCATTATCGGATGAAGGAGACGAAACAGGATGTCATTGA 2761  
Db 88217 GCTGTCTCCGTGTTCAGCGCCATTATCGGATGAAGGAGACGAAACAGGATGTCATTGA 88276  
Qy 2762 CATGCCAAAAGCTTGAAGGCTCTTGACCGGATCCATTCCTGTGAATTTTTTGCATGC 2821  
Db 88277 CATGCCAAAAGCTTGAAGGCTCTTGACCGGATCCATTCCTGTGAATTTTTTGCATGC 88336  
Qy 2822 AATTGATGGCAGCGCTTAGAAGCGCTCAACGAAATTAACCCGCTGTATTGTTTAAAGT 2881  
Db 88337 AATTGATGGCAGCGCTTAGAAGCGCTCAACGAAATTAACCCGCTGTATTGTTTAAAGT 88396  
Qy 2882 GCTGGCGTGTTCGGTTTTATCAATCCATCAAAAAGAAATTCGCATTTCCGAGGAAGAGA 2941  
Db 88397 GCTGGCGTGTTCGGTTTTATCAATCCATCAAAAAGAAATTCGCATTTCCGAGGAAGAGA 88456  
Qy 2942 GGTCAATCTCCGACATTTGACGCCATTTAGGCGTTTACGCCCAACTCCATTTTGTGCG 3001  
Db 88457 GGTCAATCTCCGACATTTGACGCCATTTAGGCGTTTACGCCCAACTCCATTTTGTGCG 88516  
Qy 3002 AGACTACTTAACTGCGCGGCAAGAGGACGAGGATCAATAAATGCTGAGTGATTT 3061  
Db 88517 AGACTACTTAACTGCGCGGCAAGAGGACGAGGATCAATAAATGCTGAGTGATTT 88576  
Qy 3062 AGGCTTTGAATTTGAATCAGTCGAAGAAATCAAGGCTAGTTTAAAGTGCAGAAAGCTGAAA 3121  
Db 88577 AGGCTTTGAATTTGAATCAGTCGAAGAAATCAAGGCTAGTTTAAAGTGCAGAAAGCTGAAA 88636  
Qy 3122 GAATCAATAAAGCAATCGGTATGATCGGAAT 3154  
Db 88637 GAATCAATAAAGCAATCGGTATGATCGGAAT 88669

RESULT 9  
AB088066  
LOCUS  
DEFINITION  
AB088066  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PUBMED

AB088066 6739 bp DNA linear BCT 17-JUL-2004  
Bacillus subtilis biotin operon (bioW, bioA, bioF, bioD, bioB, bioI), complete cds.  
AB088066  
AB088066.1 GI:21734667  
Bacillus subtilis  
Bacillus subtilis  
Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
1  
Sasaki, M., Kawamura, F. and Kurusu, Y.  
Genetic analysis of an incomplete bio operon in a biotin auxotrophic strain of Bacillus subtilis natto OK2  
J Biochem Biotechnol. Biochem. 68 (3), 739-742 (2004)  
15056910

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source

2 (bases 1 to 6739)  
Kurusu, Y., Sasaki, M. and Kotanagi, T.  
Direct Submission  
Submitted (10-JUL-2002) Yasurou Kurusu, School of Agriculture,  
Ibaraki University, Laboratory of Molecular Microbiology, 3-21-1,  
Chuo, Ami, Inashiki, Ibaraki 300-0393, Japan  
(E-mail:krsy@ipc.ibaraki.ac.jp, Tel:81-298-88-8646)  
Location/Qualifiers  
1..6739  
/organism="Bacillus subtilis"  
/mol\_type="genomic DNA"  
/isolate="OK2"  
/db\_xref="taxon:1423"  
/note="synonym:Bacillus natto"  
561..1238  
/gene="bioW"  
561..1238  
/gene="bioW"  
/codon\_start=1  
/transl\_table=11  
/product="6-carboxyhexanoate-CoA ligase"  
/protein\_id="BAC03239.1"  
/db\_xref="GI:21734668"  
/translation="MMQEEFFYSVVRMASNMGSHEDGGKHISGGERLIPHEMKHTVN  
ALLEKGLSHRGKPFDMQIQFEEVHESIKTIQPLPVYTNVSCGEGQKLARLLEKE  
GHSRVDIEKAYEQIPEWSDVRGAVLFDIHTGKMDQTKBKGVRSRMDWPDANFEKWA  
LVSHVPAHSRIKEALALASKVSRHPAAIAELCWSDDPDYITGTVAGKMGYQRTAMK  
EYCTERG"  
1330..2676  
/gene="bioA"  
1330..2676  
/gene="bioA"  
/codon\_start=1  
/transl\_table=11  
/product="adenosylmethionine-8-amino-7-oxononanoate  
aminotransferase"  
/protein\_id="BAC03240.1"  
/db\_xref="GI:21734669"  
/translation="MTHDLIEKSKHLWLPFTOMKYDENPLIIESGTGKIQVDINGK  
EYDGRSSVNVGHKKELDNAIKQIKGIAHSTLLGNTVNPATOLAETLIDISPK  
KLTRFVSDSGAEMAEIALKMAFYKWNIGKPEKQFIAMKNGYHGTIGAVSVGSIE  
LFHVYGLPMFSEYKAPIPVYRSSEGDDECCQLELEQLLKEHEEETAALESIE  
IVGASGMIYMPGEGYLAGVRELCTYDVLMIVDEATVFGRTGMFACENHVQFDLM  
AAGKGTGGYLPATVATEDIYKAFDYDENLKTFFHSGSYTNGQJGCAVALNPLAL  
FESNIEVOVAEKSKLHLFDLQHALPHVGDTRQLGFMCGAELVRSKETKEPYPADR  
RICYKSLKMRLEGLMTRPLGDIATFLPPLASTAEELTEMVAIMKQAIHEVTSLEN"  
2666..3781  
/gene="bioF"  
2666..3781  
/gene="bioF"  
/codon\_start=1  
/transl\_table=11  
/product="8-amino-7-oxononanoate synthase"  
/protein\_id="BAC03241.1"  
/db\_xref="GI:21734670"  
/translation="MKIDSWLNDRLDIAKEAGVHRNLRSMNGAPVPERNIDGENQTVW  
SSNNYLGASDRDLIDAAQTALQOFGTSSGSRLLTGNVHSEKLEKIASFFKTEAA  
LFFSSGVLANVGLSSIPEKEDVILSDLANHASIIDGCRLSKADTVVYRHTDMNDLEN  
KLNETORYQRFFIVTGVFSMDGTIVPLDQIISLAKRYHAFVVDVDAHATCVLGDGSR  
KTSEYFGVDPDIVIGLSKAVGTEGGGAAGSAVFIDFLNHAFTIFQTALPPASCAA  
AHFAFNRITSLKNGYVVGKDHTPIIPVIVIGDAKTVIFAEKLGQGGKIYAPAIRPPTV  
APGESRIRITITSDHSNMGDIDLLDKTTFHSGIKELHII"  
3778..4473  
/gene="bioD"  
3778..4473  
/gene="bioD"  
/codon\_start=1  
/transl\_table=11  
/product="dethiobiotin synthetase"  
/protein\_id="BAC03242.1"  
/db\_xref="GI:21734671"  
/translation="MRGFFVTGDTDEVGKTVISSGLAALLKEHNRDVGIVYKPFPLSGIS  
RHPHSDTSLKDMSQTSLSHEDITPPAFKAPLAPYVAGLEGKGTITNEEVLHWGRI





Qy	1562	AAAAGGAATGAGAAATAGTGAATCGAACCAATAATAATGACTAGAGAGAAGAAGAAATGAAGA	1621
Db	3999	GGCCCGCTTGCACCATACGTTGACGGAAACTTTGAGGGAAAGACTGTCCACCATGGAAGA	4058
Qy	1622	TTGTTCATGAAATTAAGGAAACGAATATTGGATAAATATGGGATGATGTTAAGGCTATTG	1681
Db	4059	GGTTTTAAGCCATTGGGGCGGATTAGAGAAAAA-----CATGA	4097
Qy	1682	GTGTTTATGGCTCTCTTGGTCGTGAGACTGATGGGCCCTATTTCGGATATTGAGATGATGT	1741
Db	4098	ATGCTTCATCGTAGAAGGTGCGGCGGTATTTCTGTGCCATTGGGAGAGGACTATTTTGGT	4157
Qy	1742	GTGTCATGTCACAGAGAAACGAGAGTTTCAGCCATGAATGGAACAACCGGTGAGTGGGAAGG	1801
Db	4158	CAGTCATGTCATAAAAGCGTTTGCAGCTTCCCATGATTATTGTGCGCGGTCTCTACCTTGG	4217
Qy	1802	TGGAAAGTGAATTTTGTATAGCGAAGAGATTCTACTAGATTATGCAATCTCAGGTGGAATCAG	1861
Db	4218	AACCATTAAATCATACCTTTTAACTGTCAAATATGCAGAAAGCAT-----	4262
Qy	1862	ATTGGCCGCTTACACATGGTCAATTTTCTCTATTTCGCGCATTTATGATTTACAGGTGGAT	1921
Db	4263	----GGGGCTTCAATTGCCGGAATTATC-----ATCAATGGAATCAGTCACT	4306
Qy	1922	ACTTAGAGAAAGTGATATCAAACTGCTAAATCGGTGAGAAGCCCAACCGTTCACGATCGGA	1981
Db	4307	CTCCTGATGAAGATGAATAAAACCAACTCTGAGATGATTGAGCGCTTATCGGTGTTCCGA	4366
Qy	1982	TTTGTGCCCTTATCGTAGAAGAGCTGTTTGAATATGCAAGCAATGCGGTAAATATTCGTG	2041
Db	4367	TTTTAGGGGTTTACGCCAAA-----GCTGCCAACGTTTACGAAGAAACCGTTCTTA	4416
Qy	2042	TGCAAGGACCGACAAACATTTCTACCATCTTGACTGTACAGGTAGCAATGSCAGGTGCCA	2101
Db	4417	CATATGTTAAGAACCATATCAATCTATCAATCTG---ATGAATCAAAATGGGGGTATGA	4473
Qy	2102	TGTTGATTGGTCTGCATCATCGCATCTGTTATACGACGAGCGCTTCGGTCTTAACTGAAG	2161
Db	4474	GAATGAATCAATGAGTGMCACTCGCAAAACGGGTGCTGGGTGGAGCAGAAGTCACTCAAG	4533
Qy	2162	CAGTTAAGCAATCAGATCTTCTTCAGGTTATGACCAATCTGTGCGCAGTTGCTAATGTCGTG	2221
Db	4534	AAGAGGCGCTTTTCAATATTAATTCATTTGCTCTGATGAAGATATTTTGTCTATTAAATGCA	4593
Qy	2222	GTCAACTTTTCGACTCTGAGAAAACCTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGAATC	2281
Db	4594	CTTTTCAATACAGAAAAACACTTTTATGGAATAAAGTAAAGCTCAATATGATATTGAATG	4653
Qy	2282	AGGAGTGGACAGAAACGACACGGATATATAGTGGATGTGTCAAAACGCATACCAATTTTGA	2341
Db	4654	CGAAATCCGGGCTCTGCCCGGAAACTGCGGCTATTGTTCA-----GTCTGCGAT	4705
Qy	2342	TTCGAAAGCGCCGATTGAGTCTTACCGGATGGTGAATAAGAGAAACGCTGCTTGAAGCGGC	2401
Db	4706	TTCGAAAGCGCCGATTGAGTCTTATCGATGGTGAATAAGAGAAACGCTGCTTGAAGCGGC	4765
Qy	2402	GAAGCGGCGCACCATCTGTAATATCGGCACATATTGTATCGTGCAGCGCAGAGGTCC	2461
Db	4766	GAAGCGGCGCACCATCTGTAATATCGGCACATATTGTATCGTGCAGCGCAGAGGGCC	4825
Qy	2462	GTCTAAACAGAGAAAGTGGATCAGGTGCTAGATGCGGTTTTCAGGAATTAAGAGAGCTATGG	2521
Db	4826	GTCTAAACAGAGAAAGTGGATCAGGTGCTAGATGCGGTTTCAGGAATTAAGAGAGCTATGG	4885
Qy	2522	ACTGAAGATTGTGCAATGTCTTGGCATGTTTGAAGCCAGAGACGGCGAGCGGCTCAAAGA	2581
Db	4886	ACTGAAGATTGTGCGTGTCTCGACTGTTTGAAGCCAGAGACGGCGAGCGGCTCAAAGA	4945
Qy	2582	TGCAGGAGTAGACCGCTATTAATCATTAATTTGAATACGTCAAGAGAAACCAATTCAAACAT	2641
Db	4946	TGCGGAGTAGACCGCTATTAATCATTAATTTGAATACGTCAAGAGAAACCAATTCAAACAT	5005
Qy	2642	CACAACCTCAATACATACGATGACAGAGTCAATACGGTTTGAATTCGAAAAAGAAATCGGG	2701

[illegible]

```
Db      1586  ACATC-----ATTATTTATTTAGTGTGCTGAAAGTGGCTTGAAGTGTGTGATGTATGT 1639
Qy      1282  GTTTTAAAGTATTGAAACCCCTTAAATTTGGTTGCACAGAAACCCCATCTGTTAAAGT 1341
Db      1640  GTTTTAAAGTATTGAAACCCCTTAAATTTGGTTGCACAGAAACCCCATCTGTTAAAGT 1699
Qy      1342  TATAAGTGACTAAACAAATAAATAAGTAGATGGGGTTTCTTTTAAATATTATGTGCTCTA 1401
Db      1700  TATAAGTGACTAAACAAATAAATAAGTAGATGGGGTTTCTTTTAAATATTATGTGCTCTA 1759
Qy      1402  ATAGTAGCAATTTATTCAGATGAAATCAAGGGTTTATGTGCAAGCAAAAAAGTGGAA 1461
Db      1760  ATAGTAGCAATTTATTCAGATGAAATCAAGGGTTTATGTGCAAGCAAAAAAGTGGAA 1819
Qy      1462  AAGTGAGACCATGTCTTAGGAAGACGAGTTATTAAATAGCTGAATAGAAAGGGTCTCTC 1521
Db      1820  AAGTGAGACCATGTCTTAGGAAGACGAGTTATTAAATAGCTGAATAGAAAGGGTCTCTC 1879
Qy      1522  CAAATATTTCTTTATTTAGAAAAGCAAATCTAAATTTATCTGAAAAGGGAATGAGATAGT 1581
Db      1880  CAAATATTTCTTTATTTAGAAAAGCAAATCTAAATTTATCTGAAAAGGGAATGAGATAGT 1939
Qy      1582  AATGACCAATTAATAGTACTAGAGAAAGAAATGAATGTTTCATGAATTAAGGAA 1641
Db      1940  AATGACCAATTAATAGTACTAGAGAAAGAAATGAATGTTTCATGAATTAAGGAA 1999
Qy      1642  CGAATATTGGATAAATATGGGATGATGTTAAAGCTATTGGTGTGTTTATGGCTCTCTGGT 1701
Db      2000  CGAATATTGGATAAATATGGGATGATGTTAAAGCTATTGGTGTGTTTATGGCTCTCTGGT 2059
Qy      1702  CGTCAGACTGATGGGCCCTATTTCGGATATTGAGATGATGTGTGTCATGTCACAGAGGAA 1761
Db      2060  CGTCAGACTGATGGGCCCTATTTCGGATATTGAGATGATGTGTGTCATGTCACAGAGGAA 2119
Qy      1762  GCAGAGTTCAGCCATGAATGAGCAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTATGC 1821
Db      2120  GCAGAGTTCAGCCATGAATGAGCAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTATGC 2179
Qy      1822  GAAGAGATTCTACTAGATTATGATCTCAGGTGGAATCAGATTGGCGCTTTACACATGGT 1881
Db      2180  GAAGAGATTCTACTAGATTATGATCTCAGGTGGAATCAGATTGGCGCTTTACACATGGT 2239
Qy      1882  CAATTTTTCTCTATTTTGGCCGATTTATGATTCAGGTGGATCTTAGAGAAAGTGTATCAA 1941
Db      2240  CAATTTTTCTCTATTTTGGCCGATTTATGATTCAGGTGGATCTTAGAGAAAGTGTATCAA 2299
Qy      1942  ACTGCTAAATCGGTAGAACCCCAACGTTCCAGATGCGATTTGTGCGCTTTATCGTAGAA 2001
Db      2300  ACTGCTAAATCGGTAGAACCCCAACGTTCCAGATGCGATTTGTGCGCTTTATCGTAGAA 2359
Qy      2002  GAGCTGTTTGAATATGACGCAAAATGGCGTAATATTCTGTGCAAGGACCGACACATTT 2061
Db      2360  GAGCTGTTTGAATATGACGCAAAATGGCGTAATATTCTGTGCAAGGACCGACACATTT 2419
Qy      2062  CTACCATCTTTGACTGTACAGSTAGCAATGGCAGGTGCCATGTTGATTTGGTCTGCATCAT 2121
Db      2420  CTACCATCTTTGACTGTACAGSTAGCAATGGCAGGTGCCATGTTGATTTGGTCTGCATCAT 2479
Qy      2122  CGCATCTGTTATACGACGAGCGCTTCGGTCTTAATCTGAGCAGGTTTAAGCAATCAGATCTT 2181
Db      2480  CGCATCTGTTATACGACGAGCGCTTCGGTCTTAATCTGAGCAGGTTTAAGCAATCAGATCTT 2539
Qy      2182  CCTTCAGSTATTGACCATCTGTGCGCAGTTGCTAATGTCGTCAACTTTCGACCTCTGAG 2241
Db      2540  CCTTCAGSTATTGACCATCTGTGCGCAGTTGCTAATGTCGTCAACTTTCGACCTCTGAG 2599
Qy      2242  AAATCTTCGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGAC 2301
Db      2600  AAATCTTCGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGAC 2659
Qy      2302  GGATATATAGTGGATGTGCAAAACGCAATACCATTTTGAAT 2342
```

```
Db      2660  GGATATATAGTGGATGTGCAAAACGCAATACCATTTTGAAT 2700
RESULT 11
LOCUS   CQ894464
DEFINITION Sequence 1 from Patent WO2004090138.
ACCESSION CQ894464
VERSION   CQ894464.1 GI:55467386
KEYWORDS .
SOURCE   synthetic construct
ORGANISM synthetic construct
          other sequences; artificial sequences.
REFERENCE 1
AUTHORS  Barg, H. and Jahn, D.
TITLE    Method for the production of vitamin B12
JOURNAL  Patent: WO 2004090138-A 1 21-OCT-2004;
          BASF Aktiengesellschaft (DE)
FEATURES
          Location/Qualifiers
            source
              1..7350
                /organism="synthetic construct"
                /mol_type="unassigned DNA"
                /db_xref="taxon:32630"
            promoter
              9..18
                /note="xylA-Promotor"
            misc_feature
              27..953
                /note="cbiX B. megaterium"
            rep_origin
              1123..1173
                /note="oriB.coli (pBR322)"
            misc_feature
              1906..2766
                /note="amp-Resistenzgen (counterclockwise)"
            rep_origin
              3611..3686
                /note="oriB.megaterium (pE194ts)"
            rep_origin
              4008..4615
                /note="repF (counterclockwise)"
            misc_feature
              6399..7169
                /note="neo-Resistenzgen"
ORIGIN
Query Match      38.08; Score 1200.4; DB 6; Length 7350;
Best Local Similarity 99.08; Pred. No. 5.8e-250;
Matches 1233; Conservative 0; Mismatches 1; Indels 12; Gaps 2;
Qy      1102  AGCTTGGCGACGAGTGCAGATCAGGGAATGAGTTTATAAAATAAAAAACACCTCAAA 1161
Db      5925  AGCTTGGCGTGCAGTGCAGATCAGGGAATGAGTTTATAAAATAAAAAACACCTCAAA 5984
Qy      1162  AGGTGCTTTTTTTTGTAGTGGTTTGAACCTGTTCTTCTTATCTTTGATACATATAGAATA 1221
Db      5985  AGGTGCTTTTTTTTGTAGTGGTTTGAACCTGTTCTTCTTATCTTTGATACATATAGAATA 6044
Qy      1222  ACCTCATTTTTTATTTTATTTTATTTAGTTGCTGAAAGGTGCGTTGAGTCTTGTATGTATGT 1281
Db      6045  ACCTC-----ATTTTTATTTTATTTAGTTGCTGAAAGGTGCGTTGAGTCTTGTATGTATGT 6098
Qy      1282  GTTTTAAAGTATTGAAACCCCTTAAATTTGGTTGCACAGAAACCCCATCTGTTAAAGT 1341
Db      6099  GTTTTAAAGTATTGAAACCCCTTAAATTTGGTTGCACAGAAACCCCATCTGTTAAAGT 6158
Qy      1342  TATAAGTGACTAAACAAATAAATAAGTAGATGGGGTTTCTTTTAAATATTATGTGCTCTA 1401
Db      6159  TATAAGTGACTAAACAAATAAATAAGTAGATGGGGTTTCTTTTAAATATTATGTGCTCTA 6218
Qy      1402  ATAGTAGCAATTTATTCAGATGAAATCAAGGGTTTATGTGGAACAAGCAAAAAAGTGGAA 1461
Db      6219  ATAGTAGCAATTTATTCAGATGAAATCAAGGGTTTATGTGGAACAAGCAAAAAAGTGGAA 6278
Qy      1462  AAGTGAGACCATG-----TGCTTAGGAAGACGAGTTATTAAATAGCTGAATAAGACGGT 1515
Db      6279  AAGTGAGACCATGAGCTTATGCTTTAGGAAGACGAGTTATTAAATAGCTGAATAAGACGGT 6338
Qy      1516  GCTCTCCAAATATCTTATTTAGAAAGCAAAATCTAAATATATCTGAAAAAGGGGAATGAGA 1575
```

Db	6339	GCTCTCAAAATATTCTTTATTTAGAAAAGCAAACTCAAAATTTATCTGAAAAGGGAATGAGA	6398
Qy	1576	ATAGTGAATGGACCAATAATAATGACTAGAGAGAAAAGAAATGAAGATTGTTCATGAAATTT	1635
Db	6399	ATAGTGAATGGACCAATAATAATGACTAGAGAGAAAAGAAATGAAGATTGTTCATGAAATTT	6458
Qy	1636	AAGGAAAGAAATATTTGGATAAATATGCGGATGATGTTAAGGCTATTTGGTGTTTATGGCTCT	1695
Db	6459	AAGGAAAGAAATATTTGGATAAATATGCGGATGATGTTAAGGCTATTTGGTGTTTATGGCTCT	6518
Qy	1696	CTTGGTCGTACAGACTGATGGGCCCTATTTCGGATATTTGAGATGATGTGTGTCATGTCAACA	1755
Db	6519	CTTGGTCGTACAGACTGATGGGCCCTATTTCGGATATTTGAGATGATGTGTGTCATGTCAACA	6578
Qy	1756	GAGGAAGCAGAGTTTCAGCCATGAAATGACAAACCGTGAGTGCAGAGGTGGAAAGTCAATTTT	1815
Db	6579	GAGGAAGCAGAGTTTCAGCCATGAAATGACAAACCGTGAGTGCAGAGGTGGAAAGTCAATTTT	6638
Qy	1816	GATAGCGAAGAGATTTCTATAGATTAATGCAATCTCAGTGGAAATCAGATTTGGCGCCTTACA	1875
Db	6639	GATAGCGAAGAGATTTCTATAGATTAATGCAATCTCAGTGGAAATCAGATTTGGCGCCTTACA	6698
Qy	1876	CATGGTCAATTTTTTCTCTATTTTTGCCGATTTATGATTTCAGGTGCGATACTTTAGAGAAAGTG	1935
Db	6699	CATGGTCAATTTTTTCTCTATTTTTGCCGATTTATGATTTCAGGTGCGATACTTTAGAGAAAGTG	6758
Qy	1936	TATCAAACTGCTAAATCGGTAGAGGCCAAACGGTTCCACATGCGATTTGTGCCCTTATC	1995
Db	6759	TATCAAACTGCTAAATCGGTAGAGGCCAAACGGTTCCACATGCGATTTGTGCCCTTATC	6818
Qy	1996	GTAGAAGAGCTGTTTGAATATGCAAGGCCAAATGGCGTAAATATTCGTGTGCAAGACCGACA	2055
Db	6819	GTAGAAGAGCTGTTTGAATATGCAAGGCCAAATGGCGTAAATATTCGTGTGCAAGACCGACA	6878
Qy	2056	ACATTTCTACCATCTTGACTGTACAGTAGCAATGCGAGTGCATGTTTGAATTTGGTCTG	2115
Db	6879	ACATTTCTACCATCTTGACTGTACAGTAGCAATGCGAGTGCATGTTTGAATTTGGTCTG	6938
Qy	2116	CATCATCGCATCTGTTATACGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCA	2175
Db	6939	CATCATCGCATCTGTTATACGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCA	6998
Qy	2176	GATCTTCTTCAGGTTATGACCATCTGTGCCAGTTTCGTAAATGTCTGGTCAACTTTCCGAC	2235
Db	6999	GATCTTCTTCAGGTTATGACCATCTGTGCCAGTTTCGTAAATGTCTGGTCAACTTTCCGAC	7058
Qy	2236	TCTGAGAAACTTTCTGGAAATCGCTAGAAATTTCTGGAAATGGGATTCAGGAGTGGACAGAA	2295
Db	7059	TCTGAGAAACTTTCTGGAAATCGCTAGAAATTTCTGGAAATGGGATTCAGGAGTGGACAGAA	7118
Qy	2296	CGACACGGATATATAGTGGATGTGTCAAAACGCATACCAATTTTGAA	2341
Db	7119	CGACACGGATATATAGTGGATGTGTCAAAACGCATACCAATTTTGAA	7164

RESULT	12
Cs1115035	
LOCUS	8152 bp DNA linear PAT 08-JUL-2005
DEFINITION	Sequence 16 from Patent WO2005056799.
Cs1115035	
ACCESSION	
VERSION	Cs1115035.1 GI:70663718
KEYWORDS	.
SOURCE	synthetic construct synthetic construct other sequences; artificial sequences. . .
ORGANISM	Nielsen,A.K. and Rasmussen,M.D. A cell with improved secretion mediated by mrga protein or homologue
REFERENCE	
AUTHORS	Patient: WO 2005056799-A 16 23-JUN-2005;
TITLE	Novozymes A/S (DK) Location/Qualifiers 1. .8152
JOURNAL	
FEATURES	
source	



AUTHORS Olsen, C. and Rasmussen, M.D.  
TITLE Method for stable gene-amplification in a bacterial host cell  
JOURNAL Patent: WO 2005042750-A 8 12-MAY-2005;  
Novozymes A/S (DK)  
FEATURES Location/Qualifiers  
source 1..13129  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="Plasmid pCLO1154"  
misc\_feature 5067  
/note="n is a, c, g, or t"  
ORIGIN  
Query Match 37.8%; Score 1192.8; DB 6; Length 13129;  
Best Local Similarity 98.9%; Pred. No. 2.4e-248;  
Matches 1224; Conservative 0; Mismatches 7; Indels 7; Gaps 2;  
Qy 1105 TTGGCAGCAGTCCAGATCAGGGAATGAGTTTATATAAATAAAAAAGCACCTGAAAAGG 1164  
Db 11149 TGGAGGAGTATGGGAGATCAGGGAATGAGTTTATAAATAAAAAAGCACCTGAAAAGG 11090  
Qy 1165 TGTCCTTTTTCATGGTTTGAACCTTGCTTCTTCTTATCTTGATACATATAGAAATAACG 1224  
Db 11089 TGTCCTTTTTCATGGTTTGAACCTTGCTTCTTCTTATCTTGATACATATAGAAATAACG 11030  
Qy 1225 TCATTTTTCATTTTATTTTTCAGTGAAGTGGTTGAAAGTGTGGTATGTATGTGTT 1284  
Db 11029 TC-----ATTTCATTTTATTTAGTTCGTAAGAGTGGTTGAAAGTGTGGTATGTATGTGTT 10976  
Qy 1285 TTAAAGTATTGAAAACCCCTTAAATTTGTCACAGAAAACCCCATCTGTAAAGTTAT 1344  
Db 10975 TTAAAGTATTGAAAACCCCTTAAATTTGTCACAGAAAACCCCATCTGTAAAGTTAT 10916  
Qy 1345 AAGTGACTAAACAAATAACTAAATAGATGGGGTTTCTTTTAAATTAATTAATGTGCTCAATA 1404  
Db 10915 AAGTGACTAAACAAATAACTAAATAGATGGGGTTTCTTTTAAATTAATTAATGTGCTCAATA 10856  
Qy 1405 GTACATTTTATTCAGATCAAAAATCAAGGTTTATGTGCAAGCAACAAAAGTCGAAAG 1464  
Db 10855 GTAGCATTTTATTCAGATCAAAAATCAAGGTTTATGTGCAAGCAACAAAAGTCGAAAG 10796  
Qy 1465 TGAGACCATG-TGCTTAGGAAGACAGATTATTAATAGCTGAATAAGACGGTGTCTCCA 1523  
Db 10795 TGAGACCATGATGCTTAGGAAGACAGATTATTAATAGCTGAATAAGACGGTGTCTCCA 10736  
Qy 1524 AATATCTTATTTAGAAAAGCAATCTAAAATTAATCTGAAAAGGGAATGAGAATAGTGAA 1583  
Db 10735 AATATCTTATTTAGAAAAGCAATCTAAAATTAATCTGAAAAGGGAATGAGAATAGTGAA 10676  
Qy 1584 TGGACCAATAATAGCTAGAGAGAAAGATGAAGATCTTCATGAAATTAAGGAACG 1643  
Db 10675 TGGACCAATAATAATGACTAGAGAGAAAGATGAAGATGTTTCATGAAATTAAGGAACG 10616  
Qy 1644 AATATTGGATAAATATGGGATGATGTTAAGGCTATTGGTGTATTATGCTCTCTTGGTCG 1703  
Db 10615 AATATTGGATAAATATGGGATGATGTTAAGGCTATTGGTGTATTATGCTCTCTTGGTCG 10556  
Qy 1704 TCAGACTGATGGGCCCTATTCCGATPATTGAGATGATGTGTCTATGTCAACAGAGGAACG 1763  
Db 10555 TCAGACTGATGGGCCCTATTCCGATPATTGAGATGATGTGTCTATGTCAACAGAGGAACG 10496  
Qy 1764 AGAGTTACGCCATGAATGGACAAACCGGTGAGTGGNAGGTGGNAGTGAATTTTGTATAGCGA 1823  
Db 10495 AGAGTTACGCCATGAATGGACAAACCGGTGAGTGGNAGGTGGNAGTGAATTTTGTATAGCGA 10436  
Qy 1824 AGAGATTTCTACTAGATTATGATCTCAGGTGGAATCAGATTGGCCGCTTACACATGTGTCA 1883  
Db 10435 AGAGATTTCTACTAGATTATGATCTCAGGTGGAATCAGATTGGCCGCTTACACATGTGTCA 10376  
Qy 1884 ATTTTCTCTATTTTGGCCGATTTATGATTCAGGTGGATACCTTAGAGAAAGTGTATCAAAC 1943  
Db 10375 ATTTTCTCTATTTTGGCCGATTTATGATTCAGGTGGATACCTTAGAGAAAGTGTATCAAAC 10316

Qy 1944 TGCTAAATCGTAGAAGCCAAACGTTTCCACGATGCGATTTGTGCCCTTATCGTAGAAGA 2003  
Db 10315 TGCTAAATCGTAGAAGCCAAACGTTTCCACGATGCGATTTGTGCCCTTATCGTAGAAGA 10256  
Qy 2004 GCTGTTTTGAATATGACAGGCAAAATGGCGTAATATATTCGTGTGCAAGACCGACAATTTCT 2063  
Db 10255 GCTGTTTTGAATATGACAGGCAAAATGGCGTAATATATTCGTGTGCAAGACCGACAATTTCT 10196  
Qy 2064 ACCATCTTTCAGTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGGTCTGCAATCATCG 2123  
Db 10195 ACCATCTTTCAGTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGGTCTGCAATCATCG 10136  
Qy 2124 CATCTGTTTATACGACGAGCGCTTCGGTCTTAACTGAAAGCAGTTAAGCAATCAGATCTTCC 2183  
Db 10135 CATCTGTTTATACGACGAGCGCTTCGGTCTTAACTGAAAGCAGTTAAGCAATCAGATCTTCC 10076  
Qy 2184 TTCAGGTTTATGACCATCTGTGCCAGTTTCGTAATGTCTGGTCAACTTTCCGACTCTGAGAA 2243  
Db 10075 TTCAGGTTTATGACCATCTGTGCCAGTTTCGTAATGTCTGGTCAACTTTCCGACTCTGAGAA 10016  
Qy 2244 ACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGATTCAGGAGTGGACAGAACGACGCG 2303  
Db 10015 ACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGATTCAGGAGTGGACAGAACGACGCG 9956  
Qy 2304 ATATATAGTGGATGTGTCAAAAACGCATACCATTTTGAA 2341  
Db 9955 ATATATAGTGGATGTGTCAAAAACGCATACCATTTTGAA 9918  
RESULT 15  
CS119875  
LOCUS CS119875 4335 bp DNA linear PAT 08-JUL-2005  
DEFINITION Sequence 137 from Patent WO2005056782.  
ACCESSION CS119875  
VERSION CS119875.1 GI:70667773  
KEYWORDS .  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1 other sequences; artificial sequences.  
AUTHORS Amin, N.S., Boston, M.G., Bott, R.R., Cervin, M.A., Concar, E.M.,  
Gustwiller, M.E., Jones, B.E., Liebeton, K., Miracle, G.S., Oh, H.,  
Poulose, A.J., Ramer, S.W., Scheibel, J.J., Weyler, W. and Whited, G.M.  
TITLE Parhydrolyase  
JOURNAL Patent: WO 2005056782-A 137 23-JUN-2005;  
GENENCOR INTERNATIONAL, INC. (US); THE PROCTER & GAMBLE COMPANY  
(US)  
FEATURES Location/Qualifiers  
source 1..4335  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/note="pAH505 plasmid"  
ORIGIN  
Query Match 27.5%; Score 867; DB 6; Length 4335;  
Best Local Similarity 100.0%; Pred. No. 1.6e-177;  
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1475 TGCTTAGGAAGACGAGTTTATTAATAGCTGAATAAGACGGTGTCTCCAAATATTTCTAT 1534  
Db 2300 TGCTTAGGAAGACGAGTTTATTAATAGCTGAATAAGACGGTGTCTCCAAATATTTCTAT 2359  
Qy 1535 TTAGAAAAGCAATCTAAAATTTATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA 1594  
Db 2360 TTAGAAAAGCAATCTAAAATTTATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA 2419  
Qy 1595 TAATGACTAGAGAGAAGATGAAGATTTCTCATGAAATTAAGGAACGATATTTGGATA 1654  
Db 2420 TAATGACTAGAGAGAAGATGAAGATTTCTCATGAAATTAAGGAACGATATTTGGATA 2479  
Qy 1655 AATATGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCTTGGTGTGCTGAGACTGATG 1714

[illegible]

Search completed: February 11, 2006, 21:06:23  
Job time : 10503 secs



***This Page Blank (uspto)***

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2006, 17:36:45 ; Search time 1217 Seconds  
(without alignments)  
17283.295 Million cell updates/sec

Title: US-10-681-086-1  
Perfect score: 3156  
Sequence: 1 ggatccacgaggttagcagc.....atcggtatgatgtcgaaattc 3156

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_21.\*

1: Geneseq1980s.\*

2: Geneseq1990s.\*

3: Geneseq2000s.\*

4: Geneseq2001as.\*

5: Geneseq2001bs.\*

6: Geneseq2002as.\*

7: Geneseq2002bs.\*

8: Geneseq2003as.\*

9: Geneseq2003bs.\*

10: Geneseq2003cs.\*

11: Geneseq2003ds.\*

12: Geneseq2004as.\*

13: Geneseq2004bs.\*

14: Geneseq2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3156	100.0	3156	6 ABL57129	AbL57129 Bacillus
2	1612.6	51.1	8478	2 AAO81792	Aao81792 B. subtilis
3	1219	38.6	3560	10 ACF36386	AcF36386 Nucleotid
4	1200.4	38.0	7350	13 ADT25766	Adt25766 Vitamin B
5	1192.8	37.8	8152	14 AEB16616	Aeb16616 Plasmid p
6	1192.8	37.8	8644	14 AEB16603	Aeb16603 Plasmid p
7	1192.8	37.8	13129	14 AEA06372	Aea06372 Novel bac
8	1081	34.3	1170	13 ADT48504	Adt48504 Bacterial
9	867	27.5	4334	14 AEA43211	Aea43211 M. smegma
10	867	27.5	5142	5 AAF29645	Aaf29645 Plasmid p
11	867	27.5	5793	6 AAS17998	Aas17998 Plasmid p
12	867	27.5	5865	12 ADL16337	Adl16337 pBPII bac
13	867	27.5	5943	6 AAS17997	Aas17997 pBPII bac
14	867	27.5	6661	10 AAD49392	Aad49392 pMOL995 p
15	867	27.5	6694	12 ADL16336	Adl16336 pBPI back
16	867	27.5	7311	14 AEA06371	Aea06371 Novel bac
17	867	27.5	7336	11 AAN91856	Aan91856 Plasmid p
18	867	27.5	8119	2 AAT33477	Aat33477 Subtilisi
19	867	27.5	8198	12 ADL16342	Adl16342 pBPI03 ex

c	20	867	27.5	8808	2 AAT32227	Aat32227 Plasmid p
c	21	867	27.5	9144	2 AAT18951	Aat18951 Plasmid p
c	22	867	27.5	9286	12 ADL16345	Adl16345 pBPI05 ex
c	23	867	27.5	10140	2 AAT32232	Aat32232 Plasmid p
c	24	867	27.5	10929	6 ABK98591	Abk98591 pEP25 vec
c	25	867	27.5	10929	9 ACD13842	Acd13842 Plasmid p
c	26	865.4	27.4	5302	3 AAA98012	Aaa98012 Expressio
c	27	865.4	27.4	5767	3 AAA98014	Aaa98014 Expressio
c	28	865.4	27.4	7334	1 AAN80664	Aan80664 Sequence
c	29	853	27.0	5609	6 AAD23867	Aad23867 pCFC717 D
c	30	853	27.0	5609	6 AAD23868	Aad23868 pEB200 DN
c	31	853	27.0	7026	6 AAD23871	Aad23871 pEB303 pl
c	32	853	27.0	7344	6 AAD23869	Aad23869 pEB301 ch
c	33	796.8	25.2	934	2 AAT18374	Aat18374 CrtB gene
c	34	796.8	25.2	2476	2 AAT18377	Aat18377 Plasmid p
c	35	768.8	24.4	1365	10 ACD02293	AcD02293 Polymorph
c	36	762.4	24.2	1200	2 AAV02474	Aav02474 Ant(4')-I
c	37	762.4	24.2	1200	2 AAV02453	Aav02453 Ant(4')-I
c	38	762.4	24.2	1200	4 AAC81670	Aac81670 E. coli/S
c	39	762.4	24.2	4775	13 ADU49640	Adu49640 Nucleotid
c	40	759	24.0	759	4 AAF83650	Aaf83650 S. aureus
c	41	758	24.0	7456	2 AAQ10686	Aaq10686 Plasmid p
c	42	701	22.2	6837	6 AAD29902	Aad29902 Plasmid p
c	43	700.2	22.2	80841	14 ADW21121	Adw21121 B. amylol
c	44	695	22.0	6169	2 AAT39284	Aat39284 Mobiliseab
c	45	521.2	16.5	1133	6 ABK73702	Abk73702 Bacillus

## ALIGNMENTS

RESULT 1

ABL57129

ID ABL57129 standard; DNA; 3156 BP.

XX ABL57129;

AC ABL57129;

XX 05-AUG-2002 (first entry)

DT 05-AUG-2002 (first entry)

XX Bacillus subtilis bioFDB gene cassette.

DE Bacillus subtilis bioFDB gene cassette.

XX Biotin; bioFDB; auxotrophy; riboflavin; fermentation; gene; ds.

KW Biotin; bioFDB; auxotrophy; riboflavin; fermentation; gene; ds.

XX Bacillus subtilis.

OS Bacillus subtilis.

XX CH2353772-A1.

PN CH2353772-A1.

XX 08-FEB-2002.

PD 08-FEB-2002.

XX 03-AUG-2001; 2001CA-02353772.

PF 03-AUG-2001; 2001CA-02353772.

XX 08-AUG-2000; 2000US-00633927.

PR 08-AUG-2000; 2000US-00633927.

XX (HOFF ) ROCHE VITAMINS AG.

PA (HOFF ) ROCHE VITAMINS AG.

XX Stebbins JW, Schlieker HW, Hohmann H, Mouncey NJ;

PI Stebbins JW, Schlieker HW, Hohmann H, Mouncey NJ;

XX WPI; 2002-424224/62.

DR WPI; 2002-424224/62.

XX Producing a target fermentation product, e.g., riboflavin, comprises

PT providing a fermentation medium and a recombinantly-produced

PT microorganism comprising a mutation that causes auxotrophic growth of the

PT microorganism.

XX Claim 19; Page 27-28/1; 42pp; English.

PS The present sequence is the bioFDB gene cassette of Bacillus subtilis.

XX The present invention provides a process for producing a target

CC fermentation product. The process uses an auxotrophic host microorganism

CC that over-produces the fermentation product. The host microorganism is

CC grown in a culture medium supplied in excess with the substrates required

CC for production of the fermentation product and in growth-limiting amounts

CC of a substrate complementing the auxotrophy. The host cell may be

CC transformed with the present polynucleotide sequence, or a biotin  
CC auxotrophy-introducing homologue of it. A preferred host microorganism is  
CC *Bacillus subtilis* strain R840 (ATCC 6838) containing multiple copies of  
CC an engineered rib operon in plasmid pRF69, and which has the ability to  
CC over-produce riboflavin. In a decoupled process with biotin as growth-  
CC limiting substrate and glucose as fermentation substrate, biomass  
CC productivity of R850:: (pRF69) Bio- was 3-fold higher than in a coupled  
CC process, and riboflavin yield on glucose was 33% higher  
XX

SQ Sequence 3156 BP; 939 A; 587 C; 789 G; 841 T; 0 U; 0 Other;

Query Match 100.0%; Score 3156; DB 6; Length 3156;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3156; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GGATCCAGAGGTTACGAGCCCTTGAAGATTGATTCCTGGTTAAACGAGCGGTTAGACAGA	60
Db	1	GGATCCAGAGGTTACGAGCCCTTGAAGATTGATTCCTGGTTAAACGAGCGGTTAGACAGA	60
Qy	61	ATGAAGAAGCCGGGTACATCGTAACCTGCGGTCAATGGATGGAGCCCGGTTCCAGAG	120
Db	61	ATGAAGAAGCCGGGTACATCGTAACCTGCGGTCAATGGATGGAGCCCGGTTCCAGAG	120
Qy	121	AGGAATATTGATGGCGAAATCAACCGGTCTGGTCTCTCAAAACAATATTATTAGGGCTCGCA	180
Db	121	AGGAATATTGATGGCGAAATCAACCGGTCTGGTCTCTCAAAACAATATTATTAGGGCTCGCA	180
Qy	181	AGCGATAGAGTTTGATCGATGCGACCCCAACAGCATTTGCGACAGGAAAGC	240
Db	181	AGCGATAGAGTTTGATCGATGCGACCCCAACAGCATTTGCGACAGGAAAGC	240
Qy	241	AGCGGTTCCAGGTTTAAACGACAGGCAATTCGGTCTGGCATGAAGCTAGAAAGAAAGATT	300
Db	241	AGCGGTTCCAGGTTTAAACGACAGGCAATTCGGTCTGGCATGAAGCTAGAAAGAAAGATT	300
Qy	301	GCCAGCTTTAAACTGACAGAGAGCGCCCTGCTGTTTTCGAGCGGTTACTTGGCCAAATGTC	360
Db	301	GCCAGCTTTAAACTGACAGAGAGCGCCCTGCTGTTTTCGAGCGGTTACTTGGCCAAATGTC	360
Qy	361	GGTGTCTTTCATCTTCCGAGAAAGGAAGATGTCATTTTAAGTGACAGCTCAATCAT	420
Db	361	GGTGTCTTTCATCTTCCGAGAAAGGAAGATGTCATTTTAAGTGACAGCTCAATCAT	420
Qy	421	GCAAGTATGATCGAGCGTCCGACTTCTAAGGCTGATACAGTTGTTTATCGCATATT	480
Db	421	GCAAGTATGATCGAGCGTCCGACTTCTAAGGCTGATACAGTTGTTTATCGGCATATT	480
Qy	481	GATATGAATGATCTTGAATAACAGCTGAATGAATAACACAGCGTTATCAGCGCGGTTTATC	540
Db	481	GATATGAATGATCTTGAATAACAGCTGAATGAATAACACAGCGTTATCAGCGCGGTTTATC	540
Qy	541	GTAACAGAGGATGATTCAGCATGGAGGCAATTCGCCCTCTTGTATCAGATCATCTCA	600
Db	541	GTAACAGAGGATGATTCAGCATGGAGGCAATTCGCCCTCTTGTATCAGATCATCTCA	600
Qy	601	CTTGGAAACGCTATCATGCTTCTGTCGTGTTGATGATGCCAGCAACAGGAGTTTG	660
Db	601	CTTGGAAACGCTATCATGCTTCTGTCGTGTTGATGATGCCAGCAACAGGAGTTTG	660
Qy	661	GGCGATTTCGGGACAAAGGAACGAGTGAATACATTTGGTGTGTTGTCGCGCATTTATCGGC	720
Db	661	GGCGATTTCGGGACAAAGGAACGAGTGAATACATTTGGTGTGTTGTCGCGCATTTATCGGC	720
Qy	721	ACCTTAAGCAAGCTGTTGGCGCGGAGGAGGTTTTCGGCAGGATCAGCGGCTTTCATC	780
Db	721	ACCTTAAGCAAGCTGTTGGCGCGGAGGAGGTTTTCGGCAGGATCAGCGGCTTTCATC	780
Qy	781	GACTTTTTCGTAACCATGCGCAGAACATTTATCTTTTCAAAACCGCTATTTCGCCAGCCAGC	840
Db	781	GACTTTTTCGTAACCATGCGCAGAACATTTATCTTTTCAAAACCGCTATTTCGCCAGCCAGC	840
Qy	841	TGTGGGCTGTCGAGGGCTTTCAACATCATTTGAAGCCAGCAGGAAACACGACAGCTT	900
Db			

Db	841	TGTGGGCTGTCACGAGGCTTTCAACATCATTTGAAGCCAGCAGGAGGAAAAACGACAGCTT	900
Qy	901	TTATTTTCTTATATCAGCATGATCAGAACGATCTGAGATATATGGTTATGTTGGTGAAA	960
Db	901	TTATTTTCTTATATCAGCATGATCAGAACGATCTGAGATATATGGTTATGTTGGTGAAA	960
Qy	961	GGAGATCACACACCGATTATTCTGTAGTCAATGGCGATGCCATAAAACCGTCTATTTT	1020
Db	961	GGAGATCACACACCGATTATTCTGTAGTCAATGGCGATGCCATAAAACCGTCTATTTT	1020
Qy	1021	GCTGAAAACTGCGAGGCAAGGGAATTTATGCTCTCTGCCATTCGGCCGCCAACCGTTGCG	1080
Db	1021	GCTGAAAACTGCGAGGCAAGGGAATTTATGCTCTCTGCCATTCGGCCGCCAACCGTTGCG	1080
Qy	1081	CCGGGTCAAAACCGGATTTCGAAGCTTCGGGACGAGTTCGAGATCAGGGAATGAGTTTATA	1140
Db	1081	CCGGGTCAAAACCGGATTTCGAAGCTTCGGGACGAGTTCGAGATCAGGGAATGAGTTTATA	1140
Qy	1141	AAATAAAAAAGCACCTGAAAAAGGTGCTTTTTTTGATGGTTTTGAACCTGTTCTTTCTT	1200
Db	1141	AAATAAAAAAGCACCTGAAAAAGGTGCTTTTTTTGATGGTTTTGAACCTGTTCTTTCTT	1200
Qy	1201	ATCTGTGATACATATAGAAATAACGTCATTTTATTTTATTTTATTTAGTCTGCTGAAAGTGCG	1260
Db	1201	ATCTGTGATACATATAGAAATAACGTCATTTTATTTTATTTTATTTAGTCTGCTGAAAGTGCG	1260
Qy	1261	TTGAAGTGTGGTATGATGTTTTTAAAGTATTTGAAAAACCTTTAAAAATTTGGTTGCAAG	1320
Db	1261	TTGAAGTGTGGTATGATGTTTTTAAAGTATTTGAAAAACCTTTAAAAATTTGGTTGCAAG	1320
Qy	1321	AAAAACCCCATCTGTTAAAGTTAAGTGACTAAACAAATAAATAAGATAGATGGGGGTTT	1380
Db	1321	AAAAACCCCATCTGTTAAAGTTAAGTGACTAAACAAATAAATAAGATAGATGGGGGTTT	1380
Qy	1381	CTTTTAAATATATGTCCTTAATAGTAGCATTTTATTCAGATGAAAAATCAAGGGTTTATG	1440
Db	1381	CTTTTAAATATATGTCCTTAATAGTAGCATTTTATTCAGATGAAAAATCAAGGGTTTATG	1440
Qy	1441	TGGACAGACAAAAGTGGAAAAGTGAGACATGCTTATAGGAGACGAGTTATTAATAG	1500
Db	1441	TGGACAGACAAAAGTGGAAAAGTGAGACATGCTTATAGGAGACGAGTTATTAATAG	1500
Qy	1501	CTGAATAAGAACGGTGTCTCCAAATATTCTTATTTAGAAAAAGCAATCTAAAAATATCT	1560
Db	1501	CTGAATAAGAACGGTGTCTCCAAATATTCTTATTTAGAAAAAGCAATCTAAAAATATCT	1560
Qy	1561	GAAGAAGGAATGAGATAGTGAATGGACCAATAATAATAGCTAGAGAGAAAGAAATGAAG	1620
Db	1561	GAAGAAGGAATGAGATAGTGAATGGACCAATAATAATAGCTAGAGAGAAAGAAATGAAG	1620
Qy	1621	ATTGTTTCATGAATTAAGGAACGAATATTCGATGATGAGGAGGCTTATTAAGGCTATT	1680
Db	1621	ATTGTTTCATGAATTAAGGAACGAATATTCGATGATGAGGAGGCTTATTAAGGCTATT	1680
Qy	1681	GGTGTTTATGGCTCTCTTGTGCTGACATGATGGGCGCTTATTCGGATATTGAGATGATG	1740
Db	1681	GGTGTTTATGGCTCTCTTGTGCTGACATGATGGGCGCTTATTCGGATATTGAGATGATG	1740
Qy	1741	TGTGTCTATGTCAACAGAGGAAGCAGAGTTCAGCCATGAATGGACAAACCGGTGAGTGAAG	1800
Db	1741	TGTGTCTATGTCAACAGAGGAAGCAGAGTTCAGCCATGAATGGACAAACCGGTGAGTGAAG	1800
Qy	1801	GTGGAAGTGAATTTGATAGCGAGAGATTTCTACTAGATTATGATCTCAGGTGGAATCA	1860
Db	1801	GTGGAAGTGAATTTGATAGCGAGAGATTTCTACTAGATTATGATCTCAGGTGGAATCA	1860
Qy	1861	GATTGGCGGCTTACACATGGTCAATTTTCTCTATTTTTCGGATTTTATGATTCAGGTGGA	1920
Db	1861	GATTGGCGGCTTACACATGGTCAATTTTCTCTATTTTTCGGATTTTATGATTCAGGTGGA	1920
Qy	1921	TACTTAGAGAAAGTGTATCAAACTGCTAAATCGGTAGAACCCCAACGTTCCACGATGCG	1980
Db	1921	TACTTAGAGAAAGTGTATCAAACTGCTAAATCGGTAGAACCCCAACGTTCCACGATGCG	1980

```
Qy 1981 ATTTGTGCCCTTATCGTAGAGAGCTGTTTGAATATGACGCAAAATGCGCGTAATATTCGT 2040
Db 1981 ATTTGTGCCCTTATCGTAGAGAGCTGTTTGAATATGACGCAAAATGCGCGTAATATTCGT 2040
Qy 2041 GTGCAAGGACCGCAACATCTTACCACTCTTGATCTGTACAGGTAGCAATGGCAGGTGCC 2100
Db 2041 GTGCAAGGACCGCAACATCTTACCACTCTTGATCTGTACAGGTAGCAATGGCAGGTGCC 2100
Qy 2101 ATGTTGATGTCGTGCAATCGCATCTGTTATACGACGAGCGCTTCGGTCTTAACTGAA 2160
Db 2101 ATGTTGATGTCGTGCAATCGCATCTGTTATACGACGAGCGCTTCGGTCTTAACTGAA 2160
Qy 2161 GCAGTTAAGCAATCAGATCTTCTTCAGGTTATGACCAATCTGTGCCAGTTTCGTAATGCT 2220
Db 2161 GCAGTTAAGCAATCAGATCTTCTTCAGGTTATGACCAATCTGTGCCAGTTTCGTAATGCT 2220
Qy 2221 GGTCAACTTTCGGACTCTGAGAACTTCTGGAATCGTGAGAAATTTCTGGAATGGGATT 2280
Db 2221 GGTCAACTTTCGGACTCTGAGAACTTCTGGAATCGTGAGAAATTTCTGGAATGGGATT 2280
Qy 2281 CAGGAGTGGACAGAAACACACGAGATATATAGTGATGTGTCAAAACGCATACCAATTTGA 2340
Db 2281 CAGGAGTGGACAGAAACACACGAGATATATAGTGATGTGTCAAAACGCATACCAATTTGA 2340
Qy 2341 ATTCGAAAGCGCCGANTGAGTCTTACCGGATGTTGAATAGGAAACGCTGCTTGAAGCG 2400
Db 2341 ATTCGAAAGCGCCGANTGAGTCTTACCGGATGTTGAATAGGAAACGCTGCTTGAAGCG 2400
Qy 2401 CGAAGCGGGCGACGATCTGAAATACGACATATTCGTCGTCGCAAGCGCGAGGTC 2460
Db 2401 CGAAGCGGGCGACGATCTGAAATACGACATATTCGTCGTCGCAAGCGCGAGGTC 2460
Qy 2461 CGTCTACAGAGAGTGCATCGAGTCGTAGATCGGTTTCAGGAAATTAAGAGACGATG 2520
Db 2461 CGTCTACAGAGAGTGCATCGAGTCGTAGATCGGTTTCAGGAAATTAAGAGACGATG 2520
Qy 2521 GACTGAAGATTTGTGCAATGCTTGGACTGTTGAAGCCAGAGCGGCAAGCGGCTCAAAG 2580
Db 2521 GACTGAAGATTTGTGCAATGCTTGGACTGTTGAAGCCAGAGCGGCAAGCGGCTCAAAG 2580
Qy 2581 ATGCAGAGTAGACCGCTTAATCATATTTGAAATACGTACAGAGAAACCATTCAAACA 2640
Db 2581 ATGCAGAGTAGACCGCTTAATCATATTTGAAATACGTACAGAGAAACCATTCAAACA 2640
Qy 2641 TCACAACCTCACATACATACGATCAGACAGTCAATACGTTGAAATCGCAAAAGATCGG 2700
Db 2641 TCACAACCTCACATACATACGATCAGACAGTCAATACGTTGAAATCGCAAAAGATCGG 2700
Qy 2701 GGCTGTCTCGTGTTCAGGCGCCATTTATCGGGATGAAAGGAGACGAAACAGGATGTCATTG 2760
Db 2701 GGCTGTCTCGTGTTCAGGCGCCATTTATCGGGATGAAAGGAGACGAAACAGGATGTCATTG 2760
Qy 2761 ACATCGCAAAAGCTTGAAGGCTTTCAGCGGATTCATTCCTGTGAAATTTTTCGATG 2820
Db 2761 ACATCGCAAAAGCTTGAAGGCTTTCAGCGGATTCATTCCTGTGAAATTTTTCGATG 2820
Qy 2821 CAATTGATGCGACCGCTTGAAGCGCTCAACGAAATTAACCCGCTGCTATTGTTTAAAG 2880
Db 2821 CAATTGATGCGACCGCTTGAAGCGCTCAACGAAATTAACCCGCTGCTATTGTTTAAAG 2880
Qy 2881 TGCTGGCGCTGTTCCGTTTTATCAATCCATCAAAAGAAATTCGATTTCCGGAGGAAGAG 2940
Db 2881 TGCTGGCGCTGTTCCGTTTTATCAATCCATCAAAAGAAATTCGATTTCCGGAGGAAGAG 2940
Qy 2941 AGGTCAATCTCCGCAATTTGACGCCATTTAGGGCTTTTACGCCGAAACTCCATTTTTCG 3000
Db 2941 AGGTCAATCTCCGCAATTTGACGCCATTTAGGGCTTTTACGCCGAAACTCCATTTTTCG 3000
Qy 3001 GAGACTACTTAACTGCGGGCAAGAGGAGCGGAGGATCAAAAATGCTGAGTGATT 3060
Db 3001 GAGACTACTTAACTGCGGGCAAGAGGAGCGGAGGATCAAAAATGCTGAGTGATT 3060
```

```
Qy 3061 TAGGCTTTGAAGTTGAATCAGTCGAAGAAATGAAGGCTAGTTTAAAGTGCAGAAAGCTGAA 3120
Db 3061 TAGGCTTTGAAGTTGAATCAGTCGAAGAAATGAAGGCTAGTTTAAAGTGCAGAAAGCTGAA 3120
Qy 3121 AGAATCAATAAAGCAATCGGTATGATGTCGAATTC 3156
Db 3121 AGAATCAATAAAGCAATCGGTATGATGTCGAATTC 3156

RESULT 2
AAQ81792
ID AAQ81792 standard; DNA; 8478 BP.
XX
XX AAQ81792;
XX AC
XX 25-MAR-2003 (revised)
DT 26-SEP-1995 (first entry)
XX
XX B. subtilis biotin operon and flanking sequences.
XX
XX Biotin operon; recombinant production; dietary additive; animal feed;
XX vitamin supplement; research reagent; ds.
XX
XX Bacillus subtilis.
XX
XX Key Location/Qualifiers
XX terminator 249..291
XX /tag= k
XX /note= "rho-independent termination site"
XX
XX promoter 324..352
XX /tag= a
XX /note= "sigma A promoter"
XX
XX misc_feature 355..387
XX /tag= b
XX /note= "potential regulatory site"
XX
XX CDS 439..1218
XX /tag= c
XX
XX CDS 1208..2554
XX /product= "biow"
XX /tag= d
XX /product= "bioA"
XX
XX CDS 2544..3713
XX /tag= e
XX /product= "bioP"
XX
XX CDS 3710..4405
XX /tag= f
XX /product= "bioD"
XX
XX CDS 4408..5415
XX /tag= g
XX /product= "bioB"
XX
XX terminator 5423..5462
XX /tag= h
XX /note= "rho-independent termination site"
XX
XX CDS 5484..6671
XX /tag= i
XX /product= "bioI"
XX
XX CDS 6748..7509
XX /tag= j
XX /label= ORF 2
XX
XX terminator 7501..7543
XX /tag= m
XX /note= "rho-independent termination site"
XX
XX CDS 7695
XX /tag= j
XX /label= ORF 3
XX
XX BP635572-A2.
XX
XX 25-JAN-1995.
XX
XX 13-JUN-1994; 94EP-00108998.
XX
XX 25-JUN-1993; 93US-00084709.
```

PR 06-MAY-1994; 94US-00239430.  
XX (HOFF ) HOFFMANN LA ROCHE & CO AG F.  
XX Bower SG, Perkins JB, Pero JG, Yocum RR;  
XX WPI; 1995-053684/08.  
XX  
XX Biotin genes, and constructs derived from Bacillus subtilis - for  
PT improved production of recombinant biotin or biotin precursor for use in  
PT e.g. dietary supplements.  
XX  
XX Example II; Fig 14; 75pp; English.  
XX  
XX AA081792 is the B. subtilis biotin operon and flanking sequences, as part  
CC of an expression vector it can be used in the recombinant production of  
CC biotin (or biotin precursor protein). The biotin can be used as a dietary  
CC additive in animal feeds, and as a vitamin supplement for human  
CC consumption. Biotin is also useful as a reagent for research, and  
CC diagnostic procedures. (Updated on 25-MAR-2003 to correct PN field.)  
XX  
XX Sequence 8478 BP; 2432 A; 1748 C; 2088 G; 2205 T; 0 U; 5 Other;  
PS  
PS Query Match 51.1%; Score 1612.6; DB 2; Length 8478;  
PS Best Local Similarity 73.0%; Pred. No. 0;  
PS Matches 2303; Conservative 0; Mismatches 624; Indels 226; Gaps 9;  
QY 2 GATCACAGGTTACGAGCCTTGAAGATTGATCTCTGGTTAAACGAGCGTTAGACAGAA 61  
DB 2524 GATCACAGGTTACGAGCCTTGAAGATTGATCTCTGGTTAAACGAGCGTTAGACAGAA 2583  
QY 62 TGAAGAAGCGCGGTACATGTAACCTCGGTCAATGGATGGAGCGCGTTCACAGAGA 121  
DB 2584 TGAAGAAGCGCGGTACATGTAACCTCGGTCAATGGATGGAGCGCGTTCACAGAGA 2643  
QY 122 GGAATATTGATGGCGAAATCAACGGTCTGGTCTCTCAACAAATATTATTAGGCTCGCAA 181  
DB 2644 GGAATATTGATGGCGAAATCAACGGTCTGGTCTCTCAACAAATATTATTAGGCTCGCAA 2703  
QY 182 CGGATAGACGTTTGATCGATGAGCCCAACAGACATTCGACCAATTTGGGACAGGAAGCA 241  
DB 2704 CGGATAGACGTTTGATCGATGAGCCCAACAGACATTCGACCAATTTGGGACAGGAAGCA 2763  
QY 242 CGGGTTACGTTTAACGACAGCGCAATTCGGTCTGCGATGAAAGCTAGAAAGAGATTG 301  
DB 2764 CGGGTTACGTTTAACGACAGCGCAATTCGGTCTGCGATGAAAGCTAGAAAGAGATTG 2823  
QY 302 CCAGCTTTAAACTGACAGAGCGGCCCTGCTGTTTTTCGAGCGGTTACTTTGGCCCAATGTGCG 361  
DB 2824 CCAGCTTTAAACTGACAGAGCGGCCCTGCTGTTTTTCGAGCGGTTACTTTGGCCCAATGTGCG 2883  
QY 362 GTGTCTTTCACTCTTGCAGAAAAGGAAGATGTCATTTAAGTGACACAGCTCAATCATG 421  
DB 2884 GTGTCTTTCACTCTTGCAGAAAAGGAAGATGTCATTTAAGTGACACAGCTCAATCATG 2943  
QY 422 CAAGTATGATCGAGCGCTCGGACCTTTCTAAGGCTGATACAGTGTGTTATCGGCATATTG 481  
DB 2944 CAAGTATGATCGAGCGCTCGGACCTTTCTAAGGCTGATACAGTGTGTTATCGGCATATTG 3003  
QY 482 ATATGAATGATCTTGAACAACAGCTGGAATGAACAACAGCGTTATCAGCGCGGTTTTATCG 541  
DB 3004 ATATGAATGATCTTGAACAACAGCTGGAATGAACAACAGCGTTATCAGCGCGGTTTTATCG 3063  
QY 542 TAAACAGCGGAGTATTACGATGATGATGACCAATCGCCCTCTTGATCAGATCATCTCAC 601  
DB 3064 TAAACAGCGGAGTATTACGATGATGATGACCAATCGCCCTCTTGATCAGATCATCTCAC 3123  
QY 602 TTGGGAAACGCTATCATCCCTTCGTGCTGTTGATGATGCCACGACAGAGTTTGG 661  
DB 3124 TTGGGAAACGCTATCATCCCTTCGTGCTGTTGATGATGCCACGACAGAGTTTGG 3183  
QY 662 CGGATTCGGGCAAGGAACGAGTGAATACTTTGGTGTGTTGTCGCGACATTTGATCGGCA 721  
DB

DB 3184 GCGATTCGGGACAAAGGAACGAGTGAATACTTTGGTGTGTTGTGCCGACATTTGTTATCGGCA 3243  
QY 722 CTTTAAAGCAAGCTGTTGGCGGGAAGAGGTTTTGGCGGAGGATCAGCGCTCTTCATCG 781  
DB 3244 CTTTAAAGCAAGCTGTTGGCGGGAAGAGGTTTTGGCGGAGGATCAGCGCTCTTCATCG 3303  
QY 782 ACTTTTTGCTGAACATGCGCCAGAACATTTATCTTTCAAAACGCTATTCCGCCAGCCAGCT 841  
DB 3304 ACTTTTTGCTGAACATGCGCCAGAACATTTATCTTTCAAAACGCTATTCCGCCAGCCAGCT 3363  
QY 842 GTGCGGCTGCTCACAGAGCTTTCAACATCATTTGAAGCCAGCAGGAGAAACGACAGCTTT 901  
DB 3364 GTGCGGCTGCTCACAGAGCTTTCAACATCATTTGAAGCCAGCAGGAGAAACGACAGCTTT 3423  
QY 902 TATTTTCTTATATCAGCATGATCAGAACCACTGCTGAAGATATGGTTATGTTGTTGAAG 961  
DB 3424 TATTTTCTTATATCAGCATGATCAGAACCACTGCTGAAGATATGGTTATGTTGTTGAAG 3483  
QY 962 GAGATCACACACCGATTATCTCTGATGCTATGCGGATGCCATATAAAACGGTCTCTATTG 1021  
DB 3484 GAGATCACACACCGATTATCTCTGATGCTATGCGGATGCCATATAAAACGGTCTCTATTG 3543  
QY 1022 CTGAAAAACTGCGGCAAGGGAATTTATGCTCTGCGCATTCGCGCGCCAAACCGTTGCGC 1081  
DB 3544 CTGAAAAACTGCGGCAAGGGAATTTATGCTCTGCGCATTCGCGCGCCAAACCGTTGCGC 3603  
QY 1082 CGGTTGAAGCCGGAATTCGAAGCTTTGGGACGAGCTCGAGATCAGGGAATGAGTTATAA 1141  
DB 3604 CGGTTGAAGCCGGAATTCGA-----ATTA 3627  
QY 1142 AATAAAAAAGCACCTCGAAAAGGTGCTTTTTTTGATGTTTGAACCTGTTCTTTCTTA 1201  
DB 3628 CAATCAGCTCTGACCAACAGTATGGTGATATTGATCATTTGCTGCAACATTTTCATTCAA 3687  
QY 1202 TCTTGATACATATAGAAATAACGTCATTTTATTTTTTATTTTTTATGTTGCTGAAAGGTGCGT 1261  
DB 3688 TCGAAAAGGAGCTGCACATCATTTGAGGGGTTTTTTTGTGACGGGAACCTGA----- 3738  
QY 1262 TGAAGTGTGGTATGATGTTTTTAAAGTATTGAABACCTTAAATTTGGTTGCACAGA 1321  
DB 3739 -----TACAGAAAGTAGGAAACCGTTATATCCACGGTCTTTG- 3776  
QY 1322 AAAACCCCATCTGTTAAAGTTATAAGTGACTAAACAAATAACTAAATAGATGGGGTTTC 1381  
DB 3777 -----CTGCCCTTATTGAAGACATAATAGACATGTCGGGGTGA 3816  
QY 1382 TTTTAAATATTATGTCCTAATAGTAGACATTTATTAGATGAAAAATAACAAGGTTTTAGT 1441  
DB 3817 TAAACCATTTTAAAGCGGGATATCGCGCCATCATCCAGATAGTACAGTTTGTCTGAA 3876  
QY 1442 GGACAGACAAAGTGGAAAAGTGAGACCATGCTTAGGAAGACGAGTTTATAATAGC 1501  
DB 3877 AGATATGTCGAGA----- 3890  
QY 1502 TGAATAAGACGGTGTCTCCAAATATTCTTATTAGAAAAGCAAAATCTAAAATTTATCTG 1561  
DB 3891 -----CCAGTCTTTCTCATGAAGACATTAACGCTTTTGCTTCAA 3930  
QY 1562 AAAAGGAATAGAAATAGTGAATGGACCAATAATAATGACTAGAGAGAAAGATGAAGA 1621  
DB 3931 GGCGCCCTTGCACCATACGTTGCGAGGAAACCTTGAGGAGAAAGACTGTCAACATCGGAAGA 3990  
QY 1622 TTGTTTCATGAATTAAGGAACGATATTGGATAAATATGGGATGATGTTAAGGCTATTG 1681  
DB 3991 GGTTTTAAAGCCATTGGGGCGGATTAGAGAAAA-----CATGAATGCTTTCATCGTAG 4043  
QY 1682 GTGTTTATGGCTCTCTTTGGTCTGACAGTCTGATGGGCCCTTATTCGGATATTGAGATGAT 1741  
DB 4044 AAGGTGCGGGGTATTCTGTGCGCATTTGGGAGAGGACTATTTGTCAGTCATGTCATAA 4103  
QY 1742 GTGTCATGTCAACAGAGGAAGCAGAGTTTCAGCCATGAATGGAACACCGGTGAGTGAAGG 1801  
DB 4104 -----AAGCGTTGACGCTTCCCATGATTATTGTCGCGCTCTCGCTTGGAAACC 4153

1802 TGAAGTGAATTTTGTAGGAGAGAGATTTCTATAGATTATGATCTCAGGTGGAATCAG 1861  
1861 TGAAGTGAATTTTGTAGGAGAGAGATTTCTATAGATTATGATCTCAGGTGGAATCAG 1861  
4154 ATTAATCATACCTTTTAACTGTCAATATATGAGAGAGATGGGCTTCCAAATCGCGGA 4213  
1862 ATTGGCGGCTTACATAGGTCAATTTTCTCTATTTTGGCGATTTATGATTCAGGTGGAT 1921  
4214 ATTAATCATACCTTTTAACTGTCAATATATGAGAGAGATGGGCTTCCAAATCGCGGA 4273  
1922 ACTTAGAGAAAGTGTATCAAACTGCTAAATCGGTAGAGAGAGAGAGAGAGAGAGAGAG 1981  
4274 ATTGAGGCTTATGCGGTGTGCGGATTTTATAGGGTTACGCCAAAGCTTGGCCAAAGTG --- 4330  
1982 TTTGTGCCCTTATCGTAGAGAGAGCTTTTGAATATGAGAGAGAGAGAGAGAGAGAGAGAG 2041  
4331 -----ACGAAGAAACGGTTCTACATATGTTAAAGACCAT- ----- 4367  
2042 TGAAGAGACGACAACTTTCTACCATCTTGTAGCTGTAGAGTAGAATGGAGGTGCCA 2101  
4368 -----TCAATCTATCATTTACTGATGAATCAAGTGGGGGTATGA 4405  
2102 TGTGTATGCTGTGATCATCGCATCTGTTATACGAGAGAGCTTGGGTCTTAACTGAAG 2161  
4406 GAATGAATCAATGATGGAATCGAGACCGGGTGTGGCTGGAGCAGAAAGTGAATGACG 4465  
2162 CAGTTAAGCAATCAGATCTTCTTCCAGGTTATGACCACTCTGTGCGAGTTGCTTAATGTCG 2221  
4466 AAGAGGCGCTTCAATATTACATTTGCTCTGATGAATATTTTGTCTATTATGACCGGG 4525  
2222 GTCAATTTCCGACTCTGAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTC 2281  
4526 CTTTTCACATCAGAAAACTTTTACGAAAAAAGTAAAGCTCAATATGATTAATGATG 4585  
2282 AGAGTGGACAGACGAGATATATAGTGTATGTCAGAGAGAGAGAGAGAGAGAGAGAG 2341  
4586 CGAAATCCGGGCTCTGCCGGAACCTGCGGCTATTGTTTCA- -----GTCGTGGAT 4637  
2342 TTCGAAGAGCCGATTTGAGTCTTACCGGATGTTGAATAGGAAACGCTGCTTGAAGCGC 2401  
4638 TTCGAAGAGCCGATTTGAGTCTTACCGGATGTTGAATAGGAAACGCTGCTTGAAGCGC 4697  
2402 GAAGCGGCGCAGATCTGAATATCGGCAATATTTGATGTCGTCGAGAGAGAGAGAGAG 2461  
4698 GAAGCGGCGCAGATCTGAATATCGGCAATATTTGATGTCGTCGAGAGAGAGAGAGAG 4757  
2462 GTCTAACAGAGAGTGAATCAGGTCTGATGTCGTTTCAAGAAATTAAGAGAGAGATGG 2521  
4758 GTCTAACAGAGAGTGAATCAGGTCTGATGTCGTTTCAAGAAATTAAGAGAGAGATGG 4817  
2522 ACTGAAGATTTGTGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2581  
4818 ACTGAAGATTTGTGATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 4877  
2582 TGAGAGAGTGAAGTGAATCAATATTTGAATATGTCGTCGTCGTCGTCGTCGTCGTCGTCG 2641  
4878 TGAGAGAGTGAAGTGAATCAATATTTGAATATGTCGTCGTCGTCGTCGTCGTCGTCGTCG 4937  
2642 CACAACCTCATACATACATGACAGAGTCAATACGTTTGAATCGCAAGAGATCGGG 2701  
4938 CACAACCTCATACATACATGACAGAGTCAATACGTTTGAATCGCAAGAGATCGGG 4997  
2702 GCTGTCTCCGTTTCCAGCGCCATTTATCGGAGTGAAGAGAGAGAGAGAGAGAGAGAGAG 2761  
4998 GCTGTCTCCGTTTCCAGCGCCATTTATCGGAGTGAAGAGAGAGAGAGAGAGAGAGAGAG 5057  
2762 CATCGCCAAAAGCTTGAAGGCTTGTGACCGGATTCATCTCTGTGATTTTGTGATGC 2821  
5058 CATCGCCAAAAGCTTGAAGGCTTGTGACCGGATTCATCTCTGTGATTTTGTGATGC 5117  
2822 AATTGATGACCGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGG 2881  
5118 AATTGATGACCGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGG 5177

QY 2882 GCTGGCGCTGTTCCGTTTATCAATCCATCAAAAGAAATTCGATTTCCGAGAGAGAGA 2941  
Db 5178 GCTGGCGCTGTTCCGTTTATCAATCCATCAAAAGAAATTCGATTTCCGAGAGAGAGA 5237  
QY 2942 GGTCAATCTCCGACATTCGAGCCATTTAGGGCTTTAGCGCGCAAACTCCATTTTGTGCG 3001  
Db 5238 GGTCAATCTCCGACATTCGAGCCATTTAGGGCTTTAGCGCGCAAACTCCATTTTGTGCG 5297  
QY 3002 AGACTACTTAAACAATCTCCGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3061  
Db 5298 AGACTACTTAAACAATCTCCGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5357  
QY 3062 AGCTTTGAAGTTGAATCAGTCAAGAAATGAAGGCTAGTTTAAAGTCGAAAGAGCTGAAA 3121  
Db 5358 AGCTTTGAAGTTGAATCAGTCAAGAAATGAAGGCTAGTTTAAAGTCGAAAGAGCTGAAA 5417  
QY 3122 GAATCAATAAAGCAATCGGTATGATGTCGAAT 3154  
Db 5418 GAATCAATAAAGCAATCGGTATGATGTCGAAT 5450  
RESULT 3  
ACF36386  
ID ACF36386 standard; DNA; 3560 BP.  
XX ACF36386;  
AC AC  
XX AC  
DT 18-DEC-2003 (first entry)  
XX  
DE Nucleotide sequence of plasmid pNMR21.  
XX  
KW Fermentation; aerobic; microorganism; respiratory chain; riboflavin;  
KW pantothenic acid; biotin; thiamine; cytochrome bd terminal oxidase;  
KW folic acid; pyridoxine; ds.  
XX  
OS Synthetic.  
XX  
PN WO2003072785-A2.  
XX  
PD 04-SEP-2003.  
XX  
PF 20-FEB-2003; 2003WO-EP001714.  
XX  
PR 27-FEB-2002; 2002EP-00004499.  
XX  
PA (HOFF ) ROCHE VITAMINS AG.  
XX  
PI Hohmann H, Mouncey NJ, Sauer U, Zamboni N;  
XX WPI; 2003-679956/64.  
XX  
DR  
XX  
PT Preparing fermentation products, e.g. riboflavin, by aerobic culture of  
microorganisms in which efficiency of the respiratory chain is increased.  
XX  
PS Claim 6; Page 11-14; 24pp; English.  
XX  
CC The invention relates to the preparation of a target fermentation product  
by culturing the appropriate aerobic microorganism in which the  
efficiency of the respiratory chain has been increased. The method is  
especially used to produce riboflavin, but also pantothenic acid; biotin;  
thiamine; folic acid; pyridoxine or an amino acid. The method provides  
higher yields of the product than cultures of unmodified microorganisms.  
CC The present sequence represents the nucleotide sequence of a plasmid  
pNMR21 comprising a polynucleotide that prevents or reduces expression of  
the cytochrome bd terminal oxidase  
XX  
SQ Sequence 3560 BP; 946 A; 681 C; 887 G; 1046 T; 0 U; 0 Other;  
Query Match 38.6%; Score 1219; DB 10; Length 3560;  
Best Local Similarity 99.5%; Pred. No. 2.6e-293;  
Matches 1235; Conservative 0; Mismatches 0; Indels 6; Gaps 1;  
QY 1102 AGCTTGGCAGCAGGTTCGAGATCAGGGAATGAGTTTATAAAATAAAAAGACCTGAAA 1161

Db 1466 AGCTTGGCAGAGGTGAGATCAGGGAATGAGTTTATAAATATAAATAAAGACACTGAAA 1525  
Qy 1162 AGGTGTCCTTTTGTGAGGTTTGAACCTGTTCTTTCTTTATCTTGTATACATATAGAAATA 1221  
Db 1526 AGGTGTCCTTTTGTGAGGTTTGAACCTGTTCTTTCTTTATCTTGTATACATATAGAAATA 1585  
Qy 1222 ACGTCAATTTTATTTTATTTTGTGCTGAAAGTGGTGGTTGAAGTGTGATGTATGT 1281  
Db 1586 ACGTC- ----ATTTTATTTTGTGCTGAAAGTGGTGGTTGAAGTGTGATGTATGT 1639  
Qy 1282 GTTTTAAAGTATTGAAAACCCCTTAAAAATTGGTTGCACAGAAAACCCCATCTGTTAAAGT 1341  
Db 1640 GTTTTAAAGTATTGAAAACCCCTTAAAAATTGGTTGCACAGAAAACCCCATCTGTTAAAGT 1699  
Qy 1342 TATAAGTGACTAAACAAATAACTAAATAGATGGGGTTTCTTTTAAATATATATGTGCTTA 1401  
Db 1700 TATAAGTGACTAAACAAATAACTAAATAGATGGGGTTTCTTTTAAATATATATGTGCTTA 1759  
Qy 1402 ATAGTAGCATTTATTCAGATGAAAATCAAGGTTTGTAGTGCAAGCAAAAAGTGGAA 1461  
Db 1760 ATAGTAGCATTTATTCAGATGAAAATCAAGGTTTGTAGTGCAAGCAAAAAGTGGAA 1819  
Qy 1462 AAGTGAGACCATGTGCTTAGGAAGACGAGTTTATTAATAGCTGAATAAGAACGGTCTCTC 1521  
Db 1820 AAGTGAGACCATGTGCTTAGGAAGACGAGTTTATTAATAGCTGAATAAGAACGGTCTCTC 1879  
Qy 1522 CAAATATCTTTATTTAGAAAAGCAAAATCTAAAAATATCTGAAAAGGGGAATGAGAAATAGTG 1581  
Db 1880 CAAATATCTTTATTTAGAAAAGCAAAATCTAAAAATATCTGAAAAGGGGAATGAGAAATAGTG 1939  
Qy 1582 AATGACCAATAAATACACTAGAGAGAAAGAAATGAGATGTTTCATGAAATTAAGGAA 1641  
Db 1940 AATGACCAATAAATACACTAGAGAGAAAGAAATGAGATGTTTCATGAAATTAAGGAA 1999  
Qy 1642 CGAATATTTGGATAAATATGGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCTTGGT 1701  
Db 2000 CGAATATTTGGATAAATATGGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCTTGGT 2059  
Qy 1702 CGTCAGACTGATGGGCCCTATTTCGGATATTCAGATGATGTTGTGTCATGTCAACAGAGGAA 1761  
Db 2060 CGTCAGACTGATGGGCCCTATTTCGGATATTCAGATGATGTTGTGTCATGTCAACAGAGGAA 2119  
Qy 1762 GCAGAGTTCCAGCATGAATGACACCGGTGAGTGGAGGTGGAAGTGAATTTTGTATAGC 1821  
Db 2120 GCAGAGTTCCAGCATGAATGACACCGGTGAGTGGAGGTGGAAGTGAATTTTGTATAGC 2179  
Qy 1822 GAAGAGATTCTACTAGATTATGCTATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGT 1881  
Db 2180 GAAGAGATTCTACTAGATTATGCTATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGT 2239  
Qy 1882 CAATTTTCTCTATTTTCCGATTTATGATTCAGGTGGATACCTTAGAGAAAGTGTATCAA 1941  
Db 2240 CAATTTTCTCTATTTTCCGATTTATGATTCAGGTGGATACCTTAGAGAAAGTGTATCAA 2299  
Qy 1942 ACTGCTAAATCGGTAGAGCCCAACGTTCCACGATGCGATTTGTCCTTATCGTAGAA 2001  
Db 2300 ACTGCTAAATCGGTAGAGCCCAACGTTCCACGATGCGATTTGTCCTTATCGTAGAA 2359  
Qy 2002 GAGCTGTTTGAATATGCAAGGCAATGGCGTAAATATTGCTGTGCAAGGACCCGACAACTTT 2061  
Db 2360 GAGCTGTTTGAATATGCAAGGCAATGGCGTAAATATTGCTGTGCAAGGACCCGACAACTTT 2419  
Qy 2062 CTACATCTTTGACTGTACAGGTAGCAATGCGAGGTGCCATGTTGATTTGCTGCAATCAT 2121  
Db 2420 CTACATCTTTGACTGTACAGGTAGCAATGCGAGGTGCCATGTTGATTTGCTGCAATCAT 2479  
Qy 2122 CGCATCTGTTATACGACGAGCGCTTCGCTTTAACTGAAGCAGTTAAGCAATCAGATCTT 2181  
Db 2480 CGCATCTGTTATACGACGAGCGCTTCGCTTTAACTGAAGCAGTTAAGCAATCAGATCTT 2539  
Qy 2182 CCTTCAGGTTATGACCATCTGTGCGAGTTCGTAATGTCTGCTCAACTTTCCGACTCTGAG 2241

Db 2540 CCTTCAGGTTATGACCATCTGTGCCAGTTCGTAATGTCGTCAACTTTCGCACTCTGAG 2599  
Qy 2242 AAACCTTCGAAATCGCTAGAGAAATTTCTGGAATGGGATTTAGGAGTGGACAGAACGACAC 2301  
Db 2600 AAACCTTCGAAATCGCTAGAGAAATTTCTGGAATGGGATTTAGGAGTGGACAGAACGACAC 2659  
Qy 2302 GGATATATAGTGGATGTGTCAAAACGCATACCATTTTGAAT 2342  
Db 2660 GGATATATAGTGGATGTGTCAAAACGCATACCATTTTGAAT 2700

RESULT 4  
ADT25766  
ID ADT25766 standard; DNA; 7350 BP.  
XX  
AC ADT25766;  
XX  
DT 13-JAN-2005 (first entry)  
XX  
DE Vitamin B12 preparation 7350bp vector.  
XX  
KW vitamin B12; transgenic; vector; ds.  
XX  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT promoter 9..18  
FT FT /\*tag= a  
FT FT /label= xyl-A promoter  
FT FT 27..953  
FT FT /\*tag= b  
FT FT /gene= "cbix B. megaterium"  
FT FT 1123..1173  
FT FT /\*tag= c  
FT FT /label= oriE.coli (pBR322)  
FT FT /direction= Counterclockwise  
FT FT 1906..2766  
FT FT /\*tag= d  
FT FT /gene= "Amp-resistance gene"  
FT FT 3611..3686  
FT FT /\*tag= e  
FT FT /label= oriB.megabacterium (pB194ts)  
FT FT /direction= Counterclockwise  
FT FT 4008..4615  
FT FT /\*tag= f  
FT FT /label= repF  
FT FT /direction= Counterclockwise  
FT FT 6399..7169  
FT FT /\*tag= g  
FT FT /gene= "Neo-resistance gene"  
XX  
PN W02004090138-A1.  
XX  
PD 21-OCT-2004.  
XX  
PF 10-MAR-2004; 2004WO-EP002430.  
XX  
PR 12-APR-2003; 2003DE-01016888.  
XX  
PA (BADI ) BASF AG.  
XX  
PI Barg H, Jahn D;  
XX  
DR WPI; 2004-757985/74.  
XX  
PT Preparation of Vitamin B12 by culturing microorganisms that have been transformed with a vector containing a cobalt chelataase gene and inducible promoter.  
XX  
PS Claim 1; SEQ ID NO 1; 57pp; German.  
XX  
CC The invention relates to a novel method for the preparation of vitamin B12. The method involves culturing a microorganism that has been



CC transformed with a vector that includes a 7350bp sequence ADT25766. The  
CC invention further relates to a vector that contains the 7350bp sequence;  
CC and genetically modified microorganisms that contain said vector. This  
CC polynucleotide sequence represents the 7350bp vector used in the  
CC preparation of vitamin B12 of the invention.

XX  
SQ Sequence 7350 BP; 2397 A; 1340 C; 1580 G; 2033 T; 0 U; 0 Other;  
Query Match 38.0%; Score 1200.4; DB 13; Length 7350;  
Best Local Similarity 99.0%; Pred. No. 1.5e-288;  
Matches 1233; Conservative 0; Mismatches 1; Indels 12; Gaps 2;

Qy 1102 AGCTGGGCGCAGGTCGAGATCAGGGAATGAGTTTATAAATAAAAAAGCACCTGAAA 1161  
Db 5925 AGCTGGGCTGCAGTGCAGATCAGGGAATGAGTTTATAAATAAAAAAGCACCTGAAA 5984  
Qy 1162 AGGTGCTCTTTTGTGATGGTTTGAACCTGTTCTTCTATCTGATACATATAGAATA 1221  
Db 5985 AGGTGCTCTTTTGTGATGGTTTGAACCTGTTCTTCTATCTGATACATATAGAATA 6044  
Qy 1222 ACGTCATTTTATTTTATTTTATTTAGTCTGCTGAAAGTGGCTTGAAGTCTTGTATGTATGT 1281  
Db 6045 ACGTC-----ATTTTATTTTATTTTATTTTATTTAGTCTGCTGAAAGTGGCTTGAAGTCTTGTATGTATGT 6098  
Qy 1282 GTTTTAAAGTATTGAAACCCCTTAAATTTGGTTCGACAGAAAACCCCATCTGTTAAAGT 1341  
Db 6099 GTTTTAAAGTATTGAAACCCCTTAAATTTGGTTCGACAGAAAACCCCATCTGTTAAAGT 6158  
Qy 1342 TATAAGTGACTAAACAATAAATAAGATAGATGGGGGTTTCTTTTAAATATATTATGTGCTTA 1401  
Db 6159 TATAAGTGACTAAACAATAAATAAGATAGATGGGGGTTTCTTTTAAATATATTATGTGCTTA 6218  
Qy 1402 ATAGTACATTTATTCAGATGAAATAATCAAGGGTTTTAGTGGACAGACAAAAGTGGAA 1461  
Db 6219 ATAGTACATTTATTCAGATGAAATAATCAAGGGTTTTAGTGGACAGACAAAAGTGGAA 6278  
Qy 1462 AAGTGAGACCATG-----TGCTTAGGAAGACGAGTTTATTAATAGCTGAATAAGAACGGT 1515  
Db 6279 AAGTGAGACCATGAGCTTATGCTTAGGAAGACGAGTTTATTAATAGCTGAATAAGAACGGT 6338  
Qy 1516 GCTCTCAAATATTTCTTTATTTAGAAAAGCAAAATCTAAATTTATCTGAAAAGGGGAATGGA 1575  
Db 6339 GCTCTCAAATATTTCTTTATTTAGAAAAGCAAAATCTAAATTTATCTGAAAAGGGGAATGGA 6398  
Qy 1576 ATAGTGAATGACCAATAATAATGACTAGAGAGAAAGATGAAGATTGTTTCATGAATTT 1635  
Db 6399 ATAGTGAATGACCAATAATAATGACTAGAGAGAAAGATGAAGATTGTTTCATGAATTT 6458  
Qy 1636 AAGGAACGAATATTGGATAAATATGGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCT 1695  
Db 6459 AAGGAACGAATATTGGATAAATATGGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCT 6518  
Qy 1696 CTTGGTCGTGAGCTGATGGCCCTATTTCGATATTCAGATGATGTTGTCATGTCACAA 1755  
Db 6519 CTTGGTCGTGAGCTGATGGCCCTATTTCGATATTCAGATGATGTTGTCATGTCACAA 6578  
Qy 1756 GAGGAACGAGTTTCACCATGATGATGACACACCGTTCAGTGGAGGGGGAAGTGAATTTT 1815  
Db 6579 GAGGAACGAGTTTCACCATGATGATGACACACCGTTCAGTGGAGGGGGAAGTGAATTTT 6638  
Qy 1816 GATAGCGAAGAGATTCTACTAGATTATGATCTCAGGTGGAATCAGATTGGCCGCTTACA 1875  
Db 6639 GATAGCGAAGAGATTCTACTAGATTATGATCTCAGGTGGAATCAGATTGGCCGCTTACA 6698  
Qy 1876 CATGGTCAATTTTCTCTATTGTCGATTTATGATTTCAAGTGGATATCTTAGAGAAAGTG 1935  
Db 6699 CATGGTCAATTTTCTCTATTGTCGATTTATGATTTCAAGTGGATATCTTAGAGAAAGTG 6758  
Qy 1936 TATCAAACTGTAATCGGTAGAGCCCAACGTTCCACGATGCGATTGTCGCTTATC 1995  
Db 6759 TATCAAACTGTAATCGGTAGAGCCCAACGTTCCACGATGCGATTGTCGCTTATC 6818  
Qy 1996 GTAGAAGAGCTGTTTGAATATGACAGGCAAAATGGCGTAAATATTCTGTGTGCAAGGACCGACA 2055

Db 6819 GTAGAAGAGCTGTTTGAATATGACAGGCAAAATGCGTAATATTCTGTGTCAAGGACCGACA 6878  
Qy 2056 ACATTTCTACCATCTCTGACTGTACAGGTAGCAATGCGAGGTGCCATGTTGATTGCTG 2115  
Db 6879 ACATTTCTACCATCTCTGACTGTACAGGTAGCAATGCGAGGTGCCATGTTGATTGCTG 6938  
Qy 2116 CATCATCGCATCTGTTATATACGAGCGCTTCCGTTCTTAACCTGAAGCAGTTAAGCAATCA 2175  
Db 6939 CATCATCGCATCTGTTATATACGAGCGCTTCCGTTCTTAACCTGAAGCAGTTAAGCAATCA 6998  
Qy 2176 GATCTTCTCTTACGTTATGACCATCTGTGCGCAGTTGCGTAATGTCGTCAACTTTCGAC 2235  
Db 6999 GATCTTCTCTTACGTTATGACCATCTGTGCGCAGTTGCGTAATGTCGTCAACTTTCGAC 7058  
Qy 2236 TCTGAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGCACAGAA 2295  
Db 7059 TCTGAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGCACAGAA 7118  
Qy 2296 CGACACGATATATAGTGGATGTGTCAAAACGCATACCATTTTGA 2341  
Db 7119 CGACACGATATATAGTGGATGTGTCAAAACGCATACCATTTTGA 7164

RESULT 5  
AEB16616  
ID AEB16616 standard; DNA; 8152 BP.  
XX AEB16616;  
AC AEB16616;  
XX 25-AUG-2005 (first entry)  
DT Plasmid pAN213ban.  
XX  
DE Oxidative stress; DNA damage; protein secretion; amyO; ds; plasmid;  
KW amyase.  
XX  
OS Bacillus subtilis.  
OS Synthetic.  
OS Unidentified.  
XX  
XX WO2005056799-A1.  
PN 23-JUN-2005.  
PD 10-DEC-2004; 2004WO-DK000859.  
PF 10-DEC-2003; 2003DK-00001824.  
PR (NOVO ) NOVOZYMES AS.  
XX  
XX Nielsen AK, Rasmussen MD;  
PI WPI; 2005-445177/45.  
DR  
XX  
XX Novel progeny cell derived from parent cell, and comprising gene encoding metallo regulated gene A MrgA protein or its functional homolog and/or DNA segment operably linked with encoding gene, useful for producing protein of interest.  
PT  
XX  
XX Example 4; SEQ ID NO 16; 49pp; English.  
PS  
XX  
XX The invention relates to a progeny cell derived from a parent cell, and comprising a gene encoding metallo regulated gene A (MrgA) protein or its functional homolog and/or DNA segment operably linked with the encoding gene, where the gene and/or DNA segment is manipulated with respect to the parent cell, or being mutated with respect to the parent cell, where the cell produces greater amounts of MrgA protein or its functional homolog than the parent cell. Also included are enhancing secretion of a protein of interest (involving expressing the protein in the cell), preparation of the progeny cell and use of the MrgA-protein or its functional homolog for enhancing production of a protein by manipulating or mutating a cell to express greater amounts of MrgA-

CC protein or its functional homologue than the non-manipulated or non-  
CC mutated cell. MrGA, a protein involved in oxidative stress response, has  
CC been demonstrated to increase the secretion of proteins when over  
CC expressed. The progeny cell is useful for producing a protein of  
CC interest, where the protein of interest is a protease, lipase, cutinase,  
CC an amylase, galactosidase, pullulanase, cellulase, glucose isomerase,  
CC protein deulfide isomerase, cyclodextrin gluconotransferase (CGR'ase),  
CC phytase, glucose oxidase, glucosyl transferase, lactase, bilirubin  
CC oxidase, xylanase, antigenic microbial or protozoan protein, bacterial  
CC protein toxin, microbial surface protein, or a viral protein. MrGA or its  
CC functional homolog is useful for enhancing production of a protein. The  
CC B. subtilis amylase gene, amyQ, was overexpressed in B. subtilis after  
CC insertion into the pel locus [placed under the control of the PconsBan  
CC promoter] from the plasmid pAN213ban. The present sequence is a the  
CC plasmid pAN213ban.

XX  
SQ Sequence 8152 BP; 2422 A; 1604 C; 1836 G; 2290 T; 0 U; 0 Other;  
Query Match 37.8%; Score 1192.8; DB 14; Length 8152;  
Best Local Similarity 98.9%; Pred. No. 1.3e-286;  
Matches 1224; Conservative 0; Mismatches 7; Indels 7; Gaps 2;  
QY 1105 TTGGCAGCAGGTGAGATCAGGGAATGAGTTTATAAATAAAGAACACCTGAAAGG 1164  
DB 4306 TGGAGGAGTATGGGAGATCAGGGAATGAGTTTATAAATAAAGAACACCTGAAAGG 4365  
QY 1165 TGTCTTTTTCATGGTTTTCGACTTGTCTTCTTATCTTGTATACATATAGAAATACG 1224  
DB 4366 TGTCTTTTTCATGGTTTTCGACTTGTCTTCTTATCTTGTATACATATAGAAATACG 4425  
QY 1225 TCATTTTATTTTATTTTAGTTGCTGAAAGGTGCGTTGAAAGTGTGGTATGTATGTGTT 1284  
DB 4426 TC-----ATTTTATTTTAGTTGCTGAAAGGTGCGTTGAAAGTGTGGTATGTATGTGTT 4479  
QY 1285 TAAAGTATTGAAACCCCTTAAATTTGTTTGCACAGAAAAACCCCATCTGTTAAAGTTAT 1344  
DB 4480 TTAAGTATTGAAACCCCTTAAATTTGTTTGCACAGAAAAACCCCATCTGTTAAAGTTAT 4539  
QY 1345 AAGTGACTTAAACAATACTAATAGATGGGGTTTCTTTTAAATATTTATGTGCTTAATA 1404  
DB 4540 AAGTGACTTAAACAATACTAATAGATGGGGTTTCTTTTAAATATTTATGTGCTTAATA 4599  
QY 1405 GTAGCATTTATTCAGATGAAAAATCAAGGGTTTTAGTGGCAAGCAAAAAAGTGAAAAAG 1464  
DB 4600 GTAGCATTTATTCAGATGAAAAATCAAGGGTTTTAGTGGCAAGCAAAAAAGTGAAAAAG 4659  
QY 1465 TGAGACCATG-TGCTTAGGAAGAGAGATTATTAATAGCTGAATTAAGAACGGTCTCTCCA 1523  
DB 4660 TGAGACCATGATGCTTAGGAAGAGAGATTATTAATAGCTGAATTAAGAACGGTCTCTCCA 4719  
QY 1524 AATATTTCTTATTTAGNAAGCAATCTTAATATTTATCTGAAAAGGGAATGGAATAGTGA 1583  
DB 4720 AATATTTCTTATTTAGNAAGCAATCTTAATATTTATCTGAAAAGGGAATGGAATAGTGA 4779  
QY 1584 TGGACCAATAATATGACTAGGAAGAAAGAAATGAAGATTGTTTCATGAAATTAAGGAACG 1643  
DB 4780 TGGACCAATAATATGACTAGGAAGAAAGAAATGAAGATTGTTTCATGAAATTAAGGAACG 4839  
QY 1644 AATATTTGGAATAATATGGGGATGATGTTAAGGCTATTGGTGTATGGCTCTCTTGTCG 1703  
DB 4840 AATATTTGGAATAATATGGGGATGATGTTAAGGCTATTGGTGTATGGCTCTCTTGTCG 4899  
QY 1704 TCAGACTGATGGGCCCTTATTCGGATATTTGAGATCATGTGTCATGTCACAGAGGAAGC 1763  
DB 4900 TCAGACTGATGGGCCCTTATTCGGATATTTGAGATCATGTGTCATGTCACAGAGGAAGC 4959  
QY 1764 AGAGTTTCAGCCATGAATGCAACACCGGTGAGTGGAAAGTGAAGTGAATTTTGTATAGCGA 1823  
DB 4960 AGAGTTTCAGCCATGAATGCAACACCGGTGAGTGGAAAGTGAAGTGAATTTTGTATAGCGA 5019  
QY 1824 AGAGATTTCTACTAGATTATGCAATCTCAGGTGGAATTCAGATTGGCGGCTTACATAGTGTCA 1883  
DB 5020 AGAGATTTCTACTAGATTATGCAATCTCAGGTGGAATTCAGATTGGCGGCTTACATAGTGTCA 5079

QY 1884 ATTTTCTCTATTTTGGCCGATTTATGATTCAGGTGGATACTTAGAGAAAGTGTATCAAAC 1943  
DB 5080 ATTTTCTCTATTTTGGCCGATTTATGATTCAGGTGGATACTTAGAGAAAGTGTATCAAAC 5139  
QY 1944 TGTAAATCGGTAGAAAGCCCAAGCGTTCCAGATGCGATTTGTGCCCTTATTCGTAGAGA 2003  
DB 5140 TGTAAATCGGTAGAAAGCCCAAGCGTTCCAGATGCGATTTGTGCCCTTATTCGTAGAGA 5199  
QY 2004 GCTGTTTGAATATGACGGCAAAATGGCGTAATATATTCGTGTGCAAGGACCGACAATTTCT 2063  
DB 5200 GCTGTTTGAATATGACGGCAAAATGGCGTAATATATTCGTGTGCAAGGACCGACAATTTCT 5259  
QY 2064 ACCATCTTCACTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTCGTCTGCAATCATCG 2123  
DB 5260 ACCATCTTCACTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTCGTCTGCAATCATCG 5319  
QY 2124 CATCTGTATACGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCC 2183  
DB 5320 CATCTGTATACGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCC 5379  
QY 2184 TTCAGGTTATGACCATCTGTGCCAGTTTCGTAATATGTCGTGTCAACTTCCGACTCTGAGAA 2243  
DB 5380 TTCAGGTTATGACCATCTGTGCCAGTTTCGTAATATGTCGTGTCAACTTCCGACTCTGAGAA 5439  
QY 2244 ACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGAGTGGACAGAACCCACACGG 2303  
DB 5440 ACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGAGTGGACAGAACCCACACGG 5499  
QY 2304 ATATATAGTGGATGTCGTCAAAAACGCATACCATTTTGA 2341  
DB 5500 ATATATAGTGGATGTCGTCAAAAACGCATACCATTTTGA 5537

RESULT 6  
AEB16603/C  
ID AEB16603 standard; DNA; 8644 BP.  
XX AC AEB16603;  
XX DT 25-AUG-2005 (first entry)  
XX DE Plasmid pDG268neo.  
XX KW Oxidative stress; DNA damage; protein secretion; metallo regulated gene;  
XX KW MrGA; ds; plasmid; pDG268neo; amyE.  
XX OS Bacillus subtilis.  
XX OS Synthetic.  
XX OS Unidentified.  
XX FN WO2005056799-A1.  
XX PD 23-JUN-2005.  
XX PF 10-DEC-2004; 2004WO-DK000859.  
XX PR 10-DEC-2003; 2003DK-00001824.  
XX PA (NOVO ) NOVOZYMES AS.  
XX PI Nielsen AK, Raemussen MD;  
XX DR WPI; 2005-445177/45.  
XX Novel progeny cell derived from parent cell, and comprising gene encoding  
PT metallo regulated gene A MrGA protein or its functional homolog and/or  
PT DNA segment operably linked with encoding gene, useful for producing  
PT protein of interest.  
XX Example 1; SEQ ID NO 3; 49pp; English.  
XX The invention relates to a progeny cell derived from a parent cell, and  
CC

comprising a gene encoding metallo regulated gene A (MrGA) protein or its functional homolog and/or DNA segment operably linked with the encoding gene, where the gene and/or DNA segment is manipulated with respect to the parent cell, or being mutated with respect to the parent cell, where the cell produces greater amounts of MrGA protein or its functional homolog than the parent cell. Also included are enhancing secretion of a protein of interest (involving expressing the protein in the cell), preparation of the progeny cell and use of the MrGA-protein or its functional homolog for enhancing production of a protein by manipulating or mutating a cell to express greater amounts of MrGA-protein or its functional homologue than the non-manipulated or non-mutated cell. MrGA, a protein involved in oxidative stress response, has been demonstrated to increase the secretion of proteins when over expressed. The progeny cell is useful for producing a protein of interest, where the protein of interest is a protease, lipase, cutinase, an amylase, galactosidase, pullulanase, cellulase, glucose isomerase, protein disulfide isomerase, cyclodextrin gluconotransferase (CGR'ase), phytase, glucose oxidase, glucosyl transferase, lactase, bilirubin oxidase, xylanase, antigenic microbial or protozoan protein, bacterial protein toxin, microbial surface protein, or a viral protein. MrGA or its functional homolog is useful for enhancing production of a protein. The present sequence is the plasmid pDG268neo which is unable to replicate in bacillus subtilis (source of the MrGA gene) but which contains a chloramphenicol resistance marker, cam, (and restriction sites for cloning) flanked by the 5' and 3' regions of the B. subtilis amyE gene. The plasmid is used for introducing an MrGA/cam marker expression cassette into the amyE locus.

Query Match 37.8%; Score 1192.8; DB 14; Length 8644;

Best Local Similarity 98.9%; Pred. No. 1.3e-286;

Matches 1224; Conservative 0; Mismatches 7; Indels 7; Gaps 2;

Qy	1105	TTGGGACGAGTCGACATCAGGGAATGAGTTTATAAAATAAAAAAGCACTGAAAGG	1164
Db	6664	TGGAGGAGTATGGGGAGATCAGGGAATGAGTTTATAAAATAAAAAAGCACTGAAAGG	6605
Qy	1165	TGCTCTTTTGTATGTTTGAACCTGCTCTTCTTATCTTGATACATATAGAAATACG	1224
Db	6604	TGCTCTTTTGTATGTTTGAACCTGCTCTTCTTATCTTGATACATATAGAAATACG	6545
Qy	1225	TCATTTTATTTTATTTAGTTGCTGAAAGGTCGTTGAAGTGTTCGTTATGATGTGTT	1284
Db	6544	TC-----ATTTTATTTTATTTAGTTGCTGAAAGGTCGTTGAAGTGTTCGTTATGATGTGTT	6491
Qy	1285	TAAAGTATTGAAACCCCTTAAATTTGCTGACAGAAAAACCCCATCTGTTAAAGTTAT	1344
Db	6490	TAAAGTATTGAAACCCCTTAAATTTGCTGACAGAAAAACCCCATCTGTTAAAGTTAT	6431
Qy	1345	AAGTGACTAAACAAATAACTAAATAGATGGGGTTCTTTTATATTATGTGTCCTAATA	1404
Db	6430	AAGTGACTAAACAAATAACTAAATAGATGGGGTTCTTTTATATTATGTGTCCTAATA	6371
Qy	1405	GTAGCATTATTCAGATGAAAAATCAAGGGTTTGTAGTGACACAAAAAGTGAAAAAG	1464
Db	6370	GTAGCATTATTCAGATGAAAAATCAAGGGTTTGTAGTGACACAAAAAGTGAAAAAG	6311
Qy	1465	TGAGACCATG-TGCTTAGGAAGACGAGTTTAAATAGCTGAATGAAGACGGTCTCTCCA	1523
Db	6310	TGAGACCATGATGCTTAGGAAGACGAGTTTAAATAGCTGAATGAAGACGGTCTCTCCA	6251
Qy	1524	AATATCTTATTTAGAAAAACAATCTAAATATCTGAAAAAGGAATGAGATAGTAA	1583
Db	6250	AATATCTTATTTAGAAAAACAATCTAAATATCTGAAAAAGGAATGAGATAGTAA	6191
Qy	1584	TGGACCAATAATATGACTAGAGAAGAAAGAAATGAAGATTGTTTCAATGAAATTAAGGAACG	1643
Db	6190	TGGACCAATAATATGACTAGAGAAGAAAGAAATGAAGATTGTTTCAATGAAATTAAGGAACG	6131
Qy	1644	AATATCGGATAATATCGGGAGATGATTAAGGCTATTTGGTGTGTTATGGCTCTCTTGGTCG	1703
Db	6130	AATATCGGATAATATCGGGAGATGATTAAGGCTATTTGGTGTGTTATGGCTCTCTTGGTCG	6071

Qy	1704	TCAGACTGATGGGCCCTATTTCGGATATTCAGATGATGTGTGTATGTCAACACAGAGGAAGC	1763
Db	6070	TCAGACTGATGGGCCCTATTTCGGATATTCAGATGATGTGTGTATGTCAACACAGAGGAAGC	6011
Qy	1764	AGAGTTCAGCCATGAATGGAACAACCGGTGAGTGAAGGTGGAAGTGAATTTTGTATAGCGA	1823
Db	6010	AGAGTTCAGCCATGAATGGAACAACCGGTGAGTGAAGGTGGAAGTGAATTTTGTATAGCGA	5951
Qy	1824	AGAGATTCTACTAGATTATGCACTCTCAGGTGGGAATCAGATTGGCCGCTTACACATGTGCA	1883
Db	5950	AGAGATTCTACTAGATTATGCACTCTCAGGTGGGAATCAGATTGGCCGCTTACACATGTGCA	5891
Qy	1884	ATTTTCTCTTATTTTGGCCGATTTATGATTCAGGTGGGATCTTAGAGAAAAGTGTATCAAAAC	1943
Db	5890	ATTTTCTCTTATTTTGGCCGATTTATGATTCAGGTGGGATCTTAGAGAAAAGTGTATCAAAAC	5831
Qy	1944	TGCTAAATCGGTAGAGCCCAAAAGCTTTCACAGATGGGATTTGTGCCCTTATCGTAGAAGA	2003
Db	5830	TGCTAAATCGGTAGAGCCCAAAAGCTTTCACAGATGGGATTTGTGCCCTTATCGTAGAAGA	5771
Qy	2004	GCTGTTTGAATATGCAAGGCAAAATGGCGTAAATATTCGTTGTCGAAGGACCGACAACTTCT	2063
Db	5770	GCTGTTTGAATATGCAAGGCAAAATGGCGTAAATATTCGTTGTCGAAGGACCGACAACTTCT	5711
Qy	2064	ACCATCTTGACTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGGTCTGCATCATCG	2123
Db	5710	ACCATCTTGACTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGGTCTGCATCATCG	5651
Qy	2124	CATCTGTTATPACGACGAGCGCTTCGGTCTTAACTGAAGCAGTGTAAAGCAATCAGATCTTCC	2183
Db	5650	CATCTGTTATPACGACGAGCGCTTCGGTCTTAACTGAAGCAGTGTAAAGCAATCAGATCTTCC	5591
Qy	2184	TTCAGGTTATGACCATCTGTCAGATTCGTAATGTCGTTCTGCTCAACTTTCGCCACTCTGAGAA	2243
Db	5590	TTCAGGTTATGACCATCTGTCAGATTCGTAATGTCGTTCTGCTCAACTTTCGCCACTCTGAGAA	5531
Qy	2244	ACTTCTGGAAATCCCTAGAGAAATTTCTGGAATGGGATTTGAGAGTGGACAGAACGACACGG	2303
Db	5530	ACTTCTGGAAATCCCTAGAGAAATTTCTGGAATGGGATTTGAGAGTGGACAGAACGACACGG	5471
Qy	2304	ATATATAGTGGATGTGTCAAAAACGATACCATTTTGA 2341	
Db	5470	ATATATAGTGGATGTGTCAAAAACGATACCATTTTGA 5433	

RESULT 7

ABA06372/c  
ID ABA06372 standard; DNA; 13129 BP.

XX AC ABA06372;

XX XX 14-JUL-2005 (first entry)

XX DE Novel bacterial host cell-related plasmid pCLO1154 DNA SeqID8.

XX XX gene expression; protein production; chromosome; genome; cell culture;  
XX KW antibiotic-resistance; ds.

XX OS Unidentified.

XX PN WO2005042750-A1.

XX PD 12-MAY-2005.

XX XX 29-OCT-2004; 2004WO-DK000750.

XX PR 31-OCT-2003; 2003DK-00001624.

XX XX (NOVO ) NOVOZYMES AS.

XX XX Olsen C, Rasmussen MD;

XX XX





(PROC ) PROCTER & GAMBLE CO.

PA Amin NS, Boston MG, Bott RR, Cervin MA, Concar EM, Gustwiller ME;  
XX Jones BE, Liebeton K, Miracle GS, Oh H, Poulouse AU, Ramer SW;  
PI Scheibel JJ, Weyler W, Whited GM;  
XX WPI; 2005-458588/46.

PT New isolated perhydrolase having preset perhydrolysis to hydrolysis  
PT ratio, useful in cleaning, bleaching and disinfecting applications e.g.  
XX paper pulp bleaching.

PS Example 6; SEQ ID NO 137; 523pp; English.

XX The invention relates to an isolated perhydrolase (I), exhibiting a  
CC perhydrolysis to hydrolysis ratio that is greater than 1. A composition  
CC containing (I) is useful for cleaning. (I) is useful in cleaning,  
CC bleaching and disinfecting applications such as paper and pulp bleaching,  
CC fabric and garment cleaning, hard surface cleaning and personal care  
CC applications (e.g. oral care, hair care, and skin care). (I) is useful in  
CC applications where peracid hydrolysis is useful, such as in the clean up  
CC of peracids. (I) enables effective cleaning, bleaching and disinfecting  
CC over broad pH and temperature ranges. The present sequence represents a  
CC M. smegmatis perhydrolase expression vector.

SQ Sequence 4334 BP; 1354 A; 760 C; 1028 G; 1192 T; 0 U; 0 Other;

Query Match 27.5%; Score 867; DB 14; Length 4334;  
Beat Local Similarity 100.0%; Pred. No. 2e-205;  
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1475 TGCTTAGGAACGAGTATTATTAAGCTGAATGAAGACGGTCTCTCAAAATATTTCTTAT 1534  
DB 2300 TGCTTAGGAACGAGTATTATTAAGCTGAATGAAGACGGTCTCTCAAAATATTTCTTAT 2359

QY 1535 TTGAAAAAGCAATCTAAATATTTATCTGAAAGGGAATGAGAATAGTGAATGGACCAATAA 1594  
DB 2360 TTGAAAAAGCAATCTAAATATTTATCTGAAAGGGAATGAGAATAGTGAATGGACCAATAA 2419

QY 1595 TAATGACTAGAGAAGAAAGATGAAGATGTTTCATGAAATTAAGGAAGATATTTGGATA 1654  
DB 2420 TAATGACTAGAGAAGAAAGATGAAGATGTTTCATGAAATTAAGGAAGATATTTGGATA 2479

QY 1655 AATATGGGGATGATCTTAAGGCTATTGGTTTATGGCTCTCTTGGTCTGAGACTGATG 1714  
DB 2480 AATATGGGGATGATCTTAAGGCTATTGGTTTATGGCTCTCTTGGTCTGAGACTGATG 2539

QY 1715 GGCCCTATTTCGGATATTGAGATGATGTGTCTCATGTCAACAGAGGAAGCAGAGTTTCAGCC 1774  
DB 2540 GGCCCTATTTCGGATATTGAGATGATGTGTCTCATGTCAACAGAGGAAGCAGAGTTTCAGCC 2599

QY 1775 ATGAATGACAAACCGGTGAGTGAAGTGAAGTGAATTTTGTATGAGGAAGATTTCTAC 1834  
DB 2600 ATGAATGACAAACCGGTGAGTGAAGTGAAGTGAATTTTGTATGAGGAAGATTTCTAC 2659

QY 1835 TAGATTATGCATCTCAGGTGGAATCAGATTTGGCGCTTACATGTCATTTTCTCTA 1894  
DB 2660 TAGATTATGCATCTCAGGTGGAATCAGATTTGGCGCTTACATGTCATTTTCTCTA 2719

QY 1895 TTTTTCGGATTTATGATTCAGGTGATATCTAGAGAAAGTGTATCAAACTGCTAAATCGG 1954  
DB 2720 TTTTTCGGATTTATGATTCAGGTGATATCTAGAGAAAGTGTATCAAACTGCTAAATCGG 2779

QY 1955 TAGAAGCCCAACGTTCCAGTGCATGATTTGGCCCTTATCGTAGAAGCTGTTTGAAT 2014  
DB 2780 TAGAAGCCCAACGTTCCAGTGCATGATTTGGCCCTTATCGTAGAAGCTGTTTGAAT 2839

QY 2015 ATGACAGCAATGGCGTAAATTTCTGTGCAAGGACCGACAACATTTCTACCATCTTGA 2074  
DB 2840 ATGACAGCAATGGCGTAAATTTCTGTGCAAGGACCGACAACATTTCTACCATCTTGA 2899

QY 2075 CTGTACAGGTAGCAATGGCAGTGCCATGTTGATTTGGTCTGTCATCATCGCATCTGTTATA 2134

DB 2900 CTGTACAGGTAGCAATGGCAGTGCCATGTTGATTTGGTCTGTCATCATCGCATCTGTTATA 2959

QY 2135 CGACGAGCGCTTCGGTCTTAACCTGAAGCAGTTAAGCAATCAGATCTTCCCTTCAGTTATG 2194  
DB 2960 CGACGAGCGCTTCGGTCTTAACCTGAAGCAGTTAAGCAATCAGATCTTCCCTTCAGTTATG 3019

QY 2195 ACCATCTGTGCCAGTTCGTAAATGTCGTCTCAACTTTTCCGACTCTTGAGAACTTCTTGAAT 2254  
DB 3020 ACCATCTGTGCCAGTTCGTAAATGTCGTCTCAACTTTTCCGACTCTTGAGAACTTCTTGAAT 3079

QY 2255 CGCTAGAGAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACAGATATATAGTGG 2314  
DB 3080 CGCTAGAGAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACAGATATATAGTGG 3139

QY 2315 ATGTGTCAAAACGCATACCATTTTGAA 2341  
DB 3140 ATGTGTCAAAACGCATACCATTTTGAA 3166

RESULT 10  
AAF29645/C  
ID AAF29645 standard; DNA; 5142 BP.  
XX AC  
XX AAF29645;  
XX  
DT 10-APR-2001 (first entry)  
XX  
DE Plasmid pL2mTFF1v1.  
XX  
KW Mouse TFF1; mTFF; trefoil protein; antiinflammatory; gastrointestinal;  
KW anti-ulcer; peptide therapy; gastrointestinal disease; acute colitis;  
KW Crohn's disease; ulcerative colitis; plasmid pL2mTFF1v1;  
KW recombinant vector; ds.  
XX  
OS Synthetic.  
XX  
PN WO200102570-A1.  
XX  
PD 11-JAN-2001.  
XX  
XX 05-JUL-2000; 2000WO-EP006343.  
PF  
XX  
PR 05-JUL-1999; 99EP-00870143.  
XX  
XX (VLAA-) VLAAms INTERUNIVERSITAIR INST BIOTECHNOG.  
XX  
XX Hans WC, Steidler L, Remaut ER;  
PI  
XX WPI; 2001-138142/14.  
XX  
XX Recombinant Lactococcus lactis for delivering a trefoil peptide useful  
PT for treating acute or chronic gastrointestinal inflammatory diseases or  
PT disorders, e.g. acute or ulcerative colitis, acute flare-ups of Crohn's  
PT disease.  
XX  
PS Claim 18; Fig 1b; 44pp; English.  
XX  
CC The present sequence is a recombinant vector which may be used to  
CC generate a recombinant Lactococcus lactis capable of delivering a trefoil  
CC peptide in vivo. The recombinant microorganism is useful for  
CC manufacturing an agent for the delivery of a trefoil peptide to the  
CC gastrointestinal tract, and for treating gastric or intestinal diseases  
CC or disorders, or lesions caused by gastric or intestinal diseases or  
CC disorders. The microorganism may also be used for preparing medicament to  
CC be used for treating gastric and /or gastrointestinal diseases or  
CC disorders, acute gastrointestinal inflammatory diseases (e.g., acute  
CC colitis, acute flare-ups of Crohn's diseases, or ulcerative colitis), and  
CC chronic and spontaneously recurring diseases of the gastrointestinal  
CC tract comprising Crohn's disease (enteritis regionalis) and ulcerative  
CC colitis (colitis ulcerosa). Disease states which can be treated by the  
CC method or compositions comprising the recombinant microorganism or  
CC trefoil peptides include disorders of and damage to the alimentary canal,  
CC including the mouth, oesophagus, stomach and large and small intestine,

CC as well as for the protection and treatment of tissues that lie outside  
CC the alimentary canal

XX Sequence 5142 BP; 1617 A; 954 C; 927 G; 1644 T; 0 U; 0 Other;

Query Match 27.5%; Score 867; DB 5; Length 5142;  
Best Local Similarity 100.0%; Pred. No. 2.1e-205;  
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1475 TGCTTAGGAACAGGATTAATAAGCTGAATAAGACGGTCTCCCAATATTTCTTAT 1534  
Db 2914 TGCTTAGGAACAGGATTAATAAGCTGAATAAGACGGTCTCCCAATATTTCTTAT 2855  
Qy 1535 TTAGAAAAGCAATCTAAAATTTATCTGAAAAGGAATGAGAATAGTGAATGACCAATAA 1594  
Db 2854 TTAGAAAAGCAATCTAAAATTTATCTGAAAAGGAATGAGAATAGTGAATGACCAATAA 2795  
Qy 1595 TAATGACTAGNAGAAAGATGAGATTTGTCATGAATTAAGAAAGCAATATTTGATA 1654  
Db 2794 TAATGACTAGNAGAAAGATGAGATTTGTCATGAATTAAGAAAGCAATATTTGATA 2735  
Qy 1655 AATATGGGATGATGTTAAGGCTATTTGGTGTATGGCTCTCTTGGTCTGAGACTGATG 1714  
Db 2734 AATATGGGATGATGTTAAGGCTATTTGGTGTATGGCTCTCTTGGTCTGAGACTGATG 2675  
Qy 1715 GGCCCTATTCGGATATTGAGATGATGTGTCTATGTCACAGAGAAAGCAGATTCAGCC 1774  
Db 2674 GGCCCTATTCGGATATTGAGATGATGTGTCTATGTCACAGAGAAAGCAGATTCAGCC 2615  
Qy 1775 ATGAATGCAACCGGTGAGTGGAAAGTGGAGTGAATTTTGATAGCGAAGATTTCTAC 1834  
Db 2614 ATGAATGCAACCGGTGAGTGGAAAGTGGAGTGAATTTTGATAGCGAAGATTTCTAC 2555  
Qy 1835 TAGAATATGATCTCAGTGGAAATCAGATTCGGCGCTTACACATGCTCAATTTTCTCTA 1894  
Db 2554 TAGAATATGATCTCAGTGGAAATCAGATTCGGCGCTTACACATGCTCAATTTTCTCTA 2495  
Qy 1895 TTTTGGCGATTTATGATTGAGTGGATATCTAGAGAAAGTGTATCAAACTGCTAAATCGG 1954  
Db 2494 TTTTGGCGATTTATGATTGAGTGGATATCTAGAGAAAGTGTATCAAACTGCTAAATCGG 2435  
Qy 1955 TAGAAGCCCAACGTTCCAGATCGGATTTGTGCGCTTTATGCTAGAGAGCTGTTGAAT 2014  
Db 2434 TAGAAGCCCAACGTTCCAGATCGGATTTGTGCGCTTTATGCTAGAGAGCTGTTGAAT 2375  
Qy 2015 ATGCAGCAATGCGTAATATTCGTGCAAGACCGCAACATTTCTACCATCCTTGA 2074  
Db 2374 ATGCAGCAATGCGTAATATTCGTGCAAGACCGCAACATTTCTACCATCCTTGA 2315  
Qy 2075 CTGTACAGGTAGCAATGCGAGGTGCCATGTTGATGTTGCTGTCATCGCATCTGTTATA 2134  
Db 2314 CTGTACAGGTAGCAATGCGAGGTGCCATGTTGATGTTGCTGTCATCGCATCTGTTATA 2255  
Qy 2135 CGACGAGCGTTCGGTCTTAACGAGGTGAAGCAATCAGATCTTCCTTCAGGTTATG 2194  
Db 2254 CGACGAGCGTTCGGTCTTAACGAGGTGAAGCAATCAGATCTTCCTTCAGGTTATG 2195  
Qy 2195 ACCATCTGCCAGTTCGTATGCTGCTCAACTTTCCGACTCTGAGAACTTCGGAAT 2254  
Db 2194 ACCATCTGCCAGTTCGTATGCTGCTCAACTTTCCGACTCTGAGAACTTCGGAAT 2135  
Qy 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACCGCATATATAGTGG 2314  
Db 2134 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACCGCATATATAGTGG 2075  
Qy 2315 ATGTGTCAAAACGCATACCAATTTTGA 2341  
Db 2074 ATGTGTCAAAACGCATACCAATTTTGA 2048

RESULT 11

AAS17998

ID AAS17998 standard; DNA; 5793 BP.

XX

AC AAS17998;

XX 12-MAR-2002 (first entry)

XX Plasmid pMOL1809.

DE pMOL1809; cyclic; circular; GalE; ds; UDP-galactose; UDP-glucose;  
XX amplification unit copy number; industrial protein production;  
XX B. subtilis; dal locus; MOL1875.

XX Synthetic.

OS Bacillus licheniformis.

OS Bacillus subtilis.

OS Staphylococcus aureus.

XX Chimeric.

XX WO200190393-A1.

XX 29-NOV-2001.

XX 23-MAY-2001; 2001WO-DK000356.

XX 24-MAY-2000; 2000DK-00000824.

XX 30-MAY-2000; 2000US-0208052P.

XX (NOVO ) NOVOZYMES AS.

XX Rasmussen MD;

XX WPI; 2002-097661/13.

XX Increasing copy number of an amplification unit in a host cell  
XX chromosome, for industrial production of polypeptides, comprises  
XX introducing an expression cassette having a gene of interest and a gene  
XX non-functional in the chromosome.

XX Example; Page 64-66; 69pp; English.

XX The invention relates to increasing the number of copies of an  
XX amplification unit integrated into a host cell chromosome, comprising  
XX introducing a nucleic acid construct consisting of an amplification unit  
XX having an expression cassette with a copy of a gene of interest, and an  
XX expressible copy of the non-functional chromosomal gene (where the host  
XX cell has the non-functional gene), where a copy of the amplification unit  
XX integrates into the host cell chromosome, the host cell becomes  
XX susceptible to an inhibitory compound endogenously produced by the host  
XX cell when cultivated in a medium comprising a precursor, and at least one  
XX copy of the amplification unit integrates into host cell chromosome,  
XX cultivating the host cell in a medium comprising the precursor, where a  
XX chromosomally integrated copy of the amplification unit is duplicated or  
XX multiplied on the host cell chromosome, selecting a host cell comprising  
XX two or more chromosomally integrated copies of the amplification unit and  
XX optionally performing one or more cycles of using the host cell selected  
XX in each new cycle, where the number of chromosomally integrated copies  
XX of the amplification unit increases with each repeat. The method is  
XX useful for increasing the copy number of a chromosomally integrated  
XX expression cassette in a microbial strain devoid of antibiotic resistance  
XX markers for use in industrial production of polypeptides in high yield.  
XX The present sequence is replication deficient plasmid used to integrate a  
XX copy of the GalE gene from B. subtilis into the dal locus of strain  
XX MOL1875 which lacks a functional GalE gene. The GalE gene converts toxic  
XX UDP-galactose to non-toxic UDP-glucose

XX Sequence 5793 BP; 1793 A; 1001 C; 1430 G; 1569 T; 0 U; 0 Other;

XX Query Match 27.5%; Score 867; DB 6; Length 5793;

XX Best Local Similarity 100.0%; Pred. No. 2.2e-205;

XX Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1475 TGCTTAGGAACAGGATTAATAAGCTGAATAAGACGGTCTCTCAATATTTCTTAT 1534

Db 968 TGCTTAGGAACAGGATTAATAAGCTGAATAAGACGGTCTCTCAATATTTCTTAT 1027



QY 1535 TTAGAAAGCAAACTCTAAAATTTATCTGAAAAGGGAATGAGAAATAGTGAATGGACCAATAA 1594  
DB 1028 TTAGAAAGCAAACTCTAAAATTTATCTGAAAAGGGAATGAGAAATAGTGAATGGACCAATAA 1087  
QY 1595 TAATGACTAGAGAGAAAGAAATGAAGATTGTTTCATGAAATTAAGAAACGAATATTGGATA 1654  
DB 1088 TAATGACTAGAGAGAAAGAAATGAAGATTGTTTCATGAAATTAAGAAACGAATATTGGATA 1147  
QY 1655 AATATGGGATGATGTTAAGGCTATTGGTGTATTTATGGCTCTCTTGGTGGTCTGAGACTGATG 1714  
DB 1148 AATATGGGATGATGTTAAGGCTATTGGTGTATTTATGGCTCTCTTGGTGGTCTGAGACTGATG 1207  
QY 1715 GGCCTATTCCGATATTGAGATGATGTGTCTCAACAGAGGAAGCAGATTTCAGCC 1774  
DB 1208 GGCCTATTCCGATATTGAGATGATGTGTCTCAACAGAGGAAGCAGATTTCAGCC 1267  
QY 1775 ATGAATGACAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTATAGCGAAGATTTCTAC 1834  
DB 1268 ATGAATGACAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTATAGCGAAGATTTCTAC 1327  
QY 1835 TAGATTATGCAATCTCAGTGGAAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 1894  
DB 1328 TAGATTATGCAATCTCAGTGGAAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 1387  
QY 1895 TTTTGCCGATTTATGATTGAGTGGATCTTATAGAGAAAGTGTATCAAACTGCTAAATCGG 1954  
DB 1388 TTTTGCCGATTTATGATTGAGTGGATCTTATAGAGAAAGTGTATCAAACTGCTAAATCGG 1447  
QY 1955 TAGAAGCCAAACGTTCCACGATGCGATTTGTGCGCTTATCGTAGAAGAGCTGTTGAAT 2014  
DB 1448 TAGAAGCCAAACGTTCCACGATGCGATTTGTGCGCTTATCGTAGAAGAGCTGTTGAAT 1507  
QY 2015 ATGCAGGCAATGGCGTAATATTGTTGTCGAAGACCGCAACATTTCTACCATCCTTGA 2074  
DB 1508 ATGCAGGCAATGGCGTAATATTGTTGTCGAAGACCGCAACATTTCTACCATCCTTGA 1567  
QY 2075 CTGTACAGTAGCAATGGCAGGTGCCATGTTGATTGGTCTGCATCATCGCATCTGTTATA 2134  
DB 1568 CTGTACAGTAGCAATGGCAGGTGCCATGTTGATTGGTCTGCATCATCGCATCTGTTATA 1627  
QY 2135 CGACAGCGCTTCGGTCTTAATCTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 2194  
DB 1628 CGACAGCGCTTCGGTCTTAATCTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 1687  
QY 2195 ACCATCTGTGCCAGTTCGTATGTTCTGTCAACTTCCGACTCTGAGAACTTCTGGAT 2254  
DB 1688 ACCATCTGTGCCAGTTCGTATGTTCTGTCAACTTCCGACTCTGAGAACTTCTGGAT 1747  
QY 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACACGGATATATAGTGG 2314  
DB 1748 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACACGGATATATAGTGG 1807  
QY 2315 ATGTGTCAAAACCGCATACCAATTTTGA 2341  
DB 1808 ATGTGTCAAAACCGCATACCAATTTTGA 1834

RESULT 12

ADL16337  
ID ADL16337 standard; DNA; 5865 BP.

XX AC ADL16337;

XX AC ADL16337;

XX AC ADL16337;

XX AC ADL16337;

XX AC ADL16337;

XX AC ADL16337;

XX AC ADL16337;

XX AC ADL16337;

PN US2004028695-A1.  
XX 12-FEB-2004.  
XX 28-MAR-2003; 2003US-00402466.  
XX 12-APR-2002; 2002US-0372152P.  
XX (PARK/) PARK S.  
XX (GIRI/) GIRI L.  
XX Park S, Giri L;  
XX WPI; 2004-168865/16.  
XX Immunogenic composition for preparing a vaccine against a lethal  
infection of Bacillus anthracis in an animal, comprises an immunizing  
amount of a recombinant B. anthracis protective antigen protein and/or  
lethal factor protein.  
XX Disclosure; SEQ ID NO 2; 143pp; English.  
XX The present invention relates to immunogenic compositions useful for  
preparing a vaccine against a lethal infection of Bacillus anthracis in  
an animal. The compositions comprise an immunizing amount of a  
recombinant B. anthracis protective antigen (rPA) protein and a  
recombinant B. anthracis lethal factor (rLF) protein. The composition and  
methods are useful in protecting against anthrax or lethal infections  
caused by B. anthracis. The present sequence is pBPII backbone vector DNA  
used in the invention.  
XX Sequence 5865 BP; 1877 A; 1014 C; 1345 G; 1629 T; 0 U; 0 Other;

Query Match 27.5%; Score 867; DB 12; Length 5865;  
Best Local Similarity 100.0%; Pred. No. 2.2e-205;  
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1475 TGCTTAGGAAGACGAGTTTAAATAGCTGAATAAGAACGGTGTCTCTCCAAATATTCCTAT 1534  
DB 4390 TGCTTAGGAAGACGAGTTTAAATAGCTGAATAAGAACGGTGTCTCTCCAAATATTCCTAT 4449  
QY 1535 TTAGAAAAGCAAACTCTAAAATTTATCTGAAAAGGGAATGAGAAATAGTGAATGGACCAATAA 1594  
DB 4450 TTAGAAAAGCAAACTCTAAAATTTATCTGAAAAGGGAATGAGAAATAGTGAATGGACCAATAA 4509  
QY 1595 TAATGACTAGAGAGAAAGAAATGAAGATTGTTTCATGAAATTAAGGAACGAATATTGGATA 1654  
DB 4510 TAATGACTAGAGAGAAAGAAATGAAGATTGTTTCATGAAATTAAGGAACGAATATTGGATA 4569  
QY 1655 AATATGGGATGATGTTAAGGCTATTGGTGTATTTATGGCTCTCTTGGTGGTCTGAGACTGATG 1714  
DB 4570 AATATGGGATGATGTTAAGGCTATTGGTGTATTTATGGCTCTCTTGGTGGTCTGAGACTGATG 4629  
QY 1715 GGCCTATTCCGATATTGAGATGATGTGTCTCAACAGAGGAAGCAGATTTCAGCC 1774  
DB 4630 GGCCTATTCCGATATTGAGATGATGTGTCTCAACAGAGGAAGCAGATTTCAGCC 4689  
QY 1775 ATGAATGGACAACCGGTGAGTGGAGTGAAGTGAATTTGTATAGCGAAGATTTCTAC 1834  
DB 4690 ATGAATGGACAACCGGTGAGTGGAGTGAAGTGAATTTGTATAGCGAAGATTTCTAC 4749  
QY 1835 TAGATTATGCAATCTCAGTGGAAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 1894  
DB 4750 TAGATTATGCAATCTCAGTGGAAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 4809  
QY 1895 TTTTGCCGATTTATGATTGAGTGGATCTTATAGAGAAAGTGTATCAAACTGCTAAATCGG 1954  
DB 4810 TTTTGCCGATTTATGATTGAGTGGATCTTATAGAGAAAGTGTATCAAACTGCTAAATCGG 4869  
QY 1955 TAGAAGCCAAACGTTCCACGATGCGATTTGTGCGCTTATCGTAGAAGAGCTGTTGAAT 2014  
DB 4870 TAGAAGCCAAACGTTCCACGATGCGATTTGTGCGCTTATCGTAGAAGAGCTGTTGAAT 4929



QY 2015 ATGAGGCAAAATGCGGTAATATTCGTGTGCAAGGACGACAAATTTCTACCACTCTTGA 2074  
Db |||||||  
QY 4930 ATGAGGCAAAATGCGGTAATATTCGTGTGCAAGGACGACAAATTTCTACCACTCTTGA 4989  
Db |||||||  
QY 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGCTGTCGTCATCGCATCTGTTATA 2134  
Db |||||||  
QY 4990 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGCTGTCGTCATCGCATCTGTTATA 5049  
Db |||||||  
QY 2135 CGACGAGCGTTCGGGCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 2194  
Db |||||||  
QY 5050 CGACGAGCGTTCGGGCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 5109  
Db |||||||  
QY 2195 ACCATCTGTCGAGTTCGTTAACTGTCGTGTCATCTTCGCGACTCTGAGAACTTCGGAAT 2254  
Db |||||||  
QY 5110 ACCATCTGTCGAGTTCGTTAACTGTCGTGTCATCTTCGCGACTCTGAGAACTTCGGAAT 5169  
Db |||||||  
QY 2255 CGCTAGAGATTTCTGGAATGCGATTCAGGAGTGGACGACGACGATATATAGTGG 2314  
Db |||||||  
QY 5170 CGCTAGAGATTTCTGGAATGCGATTCAGGAGTGGACGACGACGATATATAGTGG 5229  
Db |||||||  
QY 2315 ATGTGTCAAAACGCATACCACTTTTGAA 2341  
Db |||||||  
QY 5230 ATGTGTCAAAACGCATACCACTTTTGAA 5256  
Db |||||||

## RESULT 13

AAS17997  
ID AAS17997 standard; DNA; 5943 BP.  
XX  
AC AAS17997;  
XX  
DT 12-MAR-2002 (first entry)  
XX  
DE Plasmid pMOL1807.  
XX  
XX pMOL1807; cyclic; circular; GalE; ds; UDP-galactose; UDP-glucose;  
KW amplification unit copy number; industrial protein production;  
KW B. subtilis; dal locus; MOL1875.  
XX  
OS Synthetic.  
OS Bacillus licheniformis.  
OS Bacillus subtilis.  
OS Staphylococcus aureus.  
OS Chimeric.  
XX  
XX WO200190393-A1.  
XX  
XX 29-NOV-2001.  
XX  
XX 23-MAY-2001; 2001WO-DK000356.  
XX  
XX 24-MAY-2000; 2000DK-00000824.  
PR 30-MAY-2000; 2000US-0208052P.  
XX  
XX (NOVO ) NOVOZYMES AS.  
XX  
XX Rasmussen MD;  
XX  
XX WPI; 2002-097661/13.  
DR  
XX  
XX Increasing copy number of an amplification unit in a host cell  
PT chromosome, for industrial production of polypeptides, comprises  
PT introducing an expression cassette having a gene of interest and a gene  
PT non-functional in the chromosome.  
XX  
XX Example; Page 63-64; 69pp; English.  
PS  
XX  
XX The invention relates to increasing the number of copies of an  
CC amplification unit integrated into a host cell chromosome, comprising  
CC introducing a nucleic acid construct consisting of an amplification unit  
CC having an expression cassette with a copy of a gene of interest, and an  
CC expressible copy of the non-functional chromosomal gene (where the host  
CC cell has the non-functional gene), where a copy of the amplification unit

CC integrates into the host cell chromosome, the host cell becomes  
CC susceptible to an inhibitory compound endogenously produced by the host  
CC cell when cultivated in a medium comprising a precursor, and at least one  
CC copy of the amplification unit integrates into host cell chromosome,  
CC cultivating the host cell in a medium comprising the precursor, where a  
CC chromosomally integrated copy of the amplification unit is duplicated or  
CC multiplied on the host cell chromosome, selecting a host cell comprising  
CC two or more chromosomally integrated copies of the amplification unit and  
CC optionally performing one or more cycles of using the host cell selected  
CC in each new cycle, where the number of chromosomally integrated copies  
CC of the amplification unit increases with each repeat. The method is  
CC useful for increasing the copy number of a chromosomally integrated  
CC expression cassette in a microbial strain devoid of antibiotic resistance  
CC markers for use in industrial production of polypeptides in high yield.  
CC The present sequence is replication deficient plasmid used to integrate a  
CC copy of the GalE gene from B. subtilis into the dal locus of strain  
CC MOL1875 which lacks a functional GalE gene. The GalE gene converts toxic  
CC UDP-galactose to non-toxic UDP-glucose  
XX  
SQ Sequence 5943 BP; 1839 A; 1023 C; 1459 G; 1622 T; 0 U; 0 Other;  
Query Match 27.5%; Score 867; DB 6; Length 5943;  
Best Local Similarity 100.0%; Pred. No. 2.2e-205;  
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1475 TGCTTAGGAAGACGAGTTATTAATAGCTGAATGAAGACGGTCTCTCCAAATATTCCTAT 1534  
Db |||||||  
QY 968 TGCTTAGGAAGACGAGTTATTAATAGCTGAATGAAGACGGTCTCTCCAAATATTCCTAT 1027  
Db |||||||  
QY 1535 TTAGAAAAGCAAACTCTAAAATTTATCTGAAAAGGGAATGAGAAATAGTGAATGGACCAATAA 1594  
Db |||||||  
QY 1028 TTAGAAAAGCAAACTCTAAAATTTATCTGAAAAGGGAATGAGAAATAGTGAATGGACCAATAA 1087  
Db |||||||  
QY 1595 TAATGACTAGAGAAGAAAGAAATGAAGATTTCTCATGAAATTAAGGAACGAATTTGGATA 1654  
Db |||||||  
QY 1088 TAATGACTAGAGAAGAAAGAAATGAAGATTTCTCATGAAATTAAGGAACGAATTTGGATA 1147  
Db |||||||  
QY 1655 AATATCGGGATGATGTTAAAGGCTATTGGTGTGTTTATGGCTCTCTTGGTCTGTCAGACTGATG 1714  
Db |||||||  
QY 1148 AATATCGGGATGATGTTAAAGGCTATTGGTGTGTTTATGGCTCTCTTGGTCTGTCAGACTGATG 1207  
Db |||||||  
QY 1715 GGCCCTATTTCGGATTTAGATGATGTTGTCATGTCACAGAGGAGAGAGTTCAGCC 1774  
Db |||||||  
QY 1208 GGCCCTATTTCGGATTTAGATGATGTTGTCATGTCACAGAGGAGAGAGTTCAGCC 1267  
Db |||||||  
QY 1775 ATGAATGGACAACCGGTGAGTGGAAGTGAAGTGAATTTTGTATAGCGAAGAGATTTCTAC 1834  
Db |||||||  
QY 1268 ATGAATGGACAACCGGTGAGTGGAAGTGAAGTGAATTTTGTATAGCGAAGAGATTTCTAC 1327  
Db |||||||  
QY 1835 TAGATTATGCATCTCAGGTGGAATCAGATTTGGCCGCTTACACATGGTCAATTTTCTCTA 1894  
Db |||||||  
QY 1328 TAGATTATGCATCTCAGGTGGAATCAGATTTGGCCGCTTACACATGGTCAATTTTCTCTA 1387  
Db |||||||  
QY 1895 TTTTGGCGATTTATGATTCAGGTGGATCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 1954  
Db |||||||  
QY 1388 TTTTGGCGATTTATGATTCAGGTGGATCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 1447  
Db |||||||  
QY 1955 TAGAAGCCCAAAACGTTCCAGATGCGATTTGTGCCCTTATCGTAGAAGAGCTGTTTGAAT 2014  
Db |||||||  
QY 1448 TAGAAGCCCAAAACGTTCCAGATGCGATTTGTGCCCTTATCGTAGAAGAGCTGTTTGAAT 1507  
Db |||||||  
QY 2015 ATGACGGCAAAATGGCGTAATATTTCTGTTGCAAGGACCGACAACTTTCTACCTCTTGA 2074  
Db |||||||  
QY 1508 ATGACGGCAAAATGGCGTAATATTTCTGTTGCAAGGACCGACAACTTTCTACCTCTTGA 1567  
Db |||||||  
QY 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGCTGTCATCGCATCTGTTATA 2134  
Db |||||||  
QY 1568 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGCTGTCATCGCATCTGTTATA 1627  
Db |||||||  
QY 2135 CGACGAGCGTTCGGGCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 2194  
Db |||||||  
QY 1628 CGACGAGCGTTCGGGCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 1687  
Db |||||||

Qy 2195 ACCATCTGTCAGTTCGTAATGCTGCTCAACTTCGACTCTGAGAACTTCTGGAAT 2254  
Db 1688 ACCATCTGTCAGTTCGTAATGCTGCTCAACTTCGACTCTGAGAACTTCTGGAAT 1747  
Qy 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGACGACCGATATATAGTGG 2314  
Db 1748 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGACGACCGATATATAGTGG 1807  
Qy 2315 ATGTGTCAAAACGCATACCAATTTTGA 2341  
Db 1808 ATGTGTCAAAACGCATACCAATTTTGA 1834

## RESULT 14

AAD49392/C  
ID AAD49392 standard; DNA; 6661 BP.

XX AC AAD49392;  
XX XX

DT 07-MAR-2003 (first entry)  
XX XX

DE pMOL995 plasmid DNA.  
XX XX

KW Pectate lyase; EC 4.2.2.2; detergent; surfactant; cleaning; dishwashing;  
KW fabric stain removal; fabric whiteness maintenance; fabric softening;  
KW fabric colour appearance; fabric dye transfer inhibition; de-pilling;  
KW dental application; oral application; colour clarification; enzyme;  
KW clay stain removal; ds.  
XX XX

OS Unidentified.  
XX XX

XX Key Location/Qualifiers  
PH misc\_feature 1..1962 a

FT /note= "Encodes PUB110 plasmid"  
FT 1963..2305  
FT terminator /tag= b  
FT /note= "Encodes transcriptional terminator from amvL gene  
FT of B. licheniformis ATCC14580 and a few introduced  
FT restriction sites"  
FT 2306..3766  
FT misc\_feature /tag= c  
FT /note= "Encodes mature part of alpha-amylase"  
FT complement (3767..4075)  
FT /tag= d  
FT /note= "Encodes the promoter and signalpeptide of alpha-  
FT amylase"  
FT 4076..6661  
FT misc\_feature /tag= e  
FT /note= "Encodes PUB110 plasmid"

XX WO200292741-A2.  
PN

XX 21-NOV-2002.  
XX XX

XX 14-MAY-2002; 2002WO-DK000315.  
XX XX

XX 14-MAY-2001; 2001DK-00000755.  
PR

XX 14-MAY-2001; 2001US-0290738P.  
XX XX

XX (NOVO ) NOVOZYMES AS.  
XX XX

XX Eskelund MB, Schuelein M, Nielsen VS, Smets J;  
PI

XX WPI; 2003-059005/05.  
XX XX

XX New polypeptide useful as ingredient of detergent composition for  
PT cleaning fabric, dishware or hard surface, encoded by DNA sequence  
PT endogenous to strain of Bacillus subtilis, comprises pectate lyase  
PT activity.  
XX XX

XX Disclosure; Page 84-87; 95pp; English.  
PS

XX

CC The present invention relates to pectate lyase (EC 4.2.2.2) proteins and  
CC polynucleotides encoding such proteins. The invention also relates to  
CC detergent compositions comprising a surfactant and pectate lyase enzyme.  
CC The detergent is useful for cleaning a fabric, a dishware or hard surface  
CC to provide superior cleaning performance, for fabric cleaning and/or  
CC fabric stain removal and/or fabric whiteness maintenance and/or fabric  
CC softening and/or fabric colour appearance and/or fabric dye transfer  
CC inhibition, for cleaning hard surfaces such as floors, walls or bathroom  
CC tiles, for hand and machine dishwashing and for oral and/or dental  
CC applications. They are useful for colour clarification, de-pilling and in  
CC clay stain removal. The present sequence is pMOL995 plasmid DNA  
XX XX

SQ Sequence 6661 BP; 1859 A; 1534 C; 991 G; 2277 T; 0 U; 0 Other;

Query Match 27.5%; Score 867; DB 10; Length 6661;

Best Local Similarity 100.0%; Pred. No. 2.3e-205;

Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1475 TGCTTAGGAAGACGAGTTATTATAATAGCTGAATAGAACGGTGTCTCCAAATATTTCTTAT 1534

Db 5180 TGCTTAGGAAGACGAGTTATTATAATAGCTGAATAGAACGGTGTCTCCAAATATTTCTTAT 5121

Qy 1535 TTAGAAAAGCAATCTAAAAATTTATCTGAAAAGGGAATGAGAATAGTAAGGACCAATAA 1594

Db 5120 TTAGAAAAGCAATCTAAAAATTTATCTGAAAAGGGAATGAGAATAGTAAGGACCAATAA 5061

Qy 1595 TAATGACTAGAGAAGAAGAAATGAAGATTGTTCTATGAAATTAAGGACGAATATGGATA 1654

Db 5060 TAATGACTAGAGAAGAAGAAATGAAGATTGTTCTATGAAATTAAGGACGAATATGGATA 5001

Qy 1655 AATATGGGATGATGTTAAGGCTATTGGTGTATTATGCTCTCTTGGTGTCTCAGACTGTATG 1714

Db 5000 AATATGGGATGATGTTAAGGCTATTGGTGTATTATGCTCTCTTGGTGTCTCAGACTGTATG 4941

Qy 1715 GGCCCTATTCGGATATTGAGATGATGTGTCTATGTCAACAGAGGAAGAGAGTTACGCC 1774

Db 4940 GGCCCTATTCGGATATTGAGATGATGTGTCTATGTCAACAGAGGAAGAGAGTTACGCC 4881

Qy 1775 ATGAATGGACAACCGGTGAGTGGAAAGTGAAGTGAATTTTGTATAGCGAGAGATTTCTAC 1834

Db 4880 ATGAATGGACAACCGGTGAGTGGAAAGTGAAGTGAATTTTGTATAGCGAGAGATTTCTAC 4821

Qy 1835 TAGATATGTCATCTCAGGTGGAATCAGATTGGCGCTTACACATGGTCAATTTTCTCTA 1894

Db 4820 TAGATATGTCATCTCAGGTGGAATCAGATTGGCGCTTACACATGGTCAATTTTCTCTA 4761

Qy 1895 TTTTGGCGATTTATGATTCAGGTGGATPACTTAGAGAAAGTGTATCAAACTGCTAAATCGG 1954

Db 4760 TTTTGGCGATTTATGATTCAGGTGGATPACTTAGAGAAAGTGTATCAAACTGCTAAATCGG 4701

Qy 1955 TAGAAGCCCAACGTTTCCAGATGCGATTTGTGCCCTTATCGTAGAGAGCTGTTGAAT 2014

Db 4700 TAGAAGCCCAACGTTTCCAGATGCGATTTGTGCCCTTATCGTAGAGAGCTGTTGAAT 4641

Qy 2015 ATGACGGCAATGCGTGAATATTTCGTGTGCAAGACCGCAACAATTTTACCCTCTTGA 2074

Db 4640 ATGACGGCAATGCGTGAATATTTCGTGTGCAAGACCGCAACAATTTTACCCTCTTGA 4581

Qy 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGCTGTGCATCATCGCATCTGTTATA 2134

Db 4580 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGCTGTGCATCATCGCATCTGTTATA 4521

Qy 2135 CGACGAGCGTTCGGTCTTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTATG 2194

Db 4520 CGACGAGCGTTCGGTCTTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTATG 4461

Qy 2195 ACCATCTGCCAGTTCGTAATGCTGTGGTCAACTTTCCGACTCTGAGAACTTCTGGAAT 2254

Db 4460 ACCATCTGCCAGTTCGTAATGCTGTGGTCAACTTTCCGACTCTGAGAACTTCTGGAAT 4401

Qy 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGACGGAATATATAGTGG 2314

Db 4400 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGACGGAATATATAGTGG 4341

QY 2315 ATGTGTCAAAACGCATACCACTTTTGAA 2341  
ADL16336  
Db 4340 ATGTGTCAAAACGCATACCACTTTTGAA 4314

RESULT 15  
ADL16336  
ID ADL16336 standard; DNA; 6694 BP.  
XX ADL16336;  
AC  
XX  
DT 06-MAY-2004 (first entry)  
XX  
XX pBPI backbone vector DNA.  
XX  
XX Immunogenic; vaccine; lethal infection; protective antigen; PA;  
KW lethal factor; LF; anthrax; antibacterial; ds.  
XX  
XX Unidentified.  
XX OS  
XX US2004028695-A1.  
XX  
XX 12-FEB-2004.  
XX  
XX 28-MAR-2003; 2003US-00402466.  
XX  
XX 12-APR-2002; 2002US-0372152P.  
XX  
XX (PARK/) PARK S.  
PA (GIRI/) GIRI L.  
XX  
XX Park S, Giri L;  
PI  
XX WPI; 2004-168865/16.  
XX  
XX Immunogenic composition for preparing a vaccine against a lethal  
PT infection of Bacillus anthracis in an animal, comprises an immunizing  
PT amount of a recombinant B. anthracis protective antigen protein and/or  
PT lethal factor protein.  
XX  
XX Disclosure; SEQ ID NO 1; 143pp; English.  
XX  
XX The present invention relates to immunogenic compositions useful for  
CC preparing a vaccine against a lethal infection of Bacillus anthracis in  
CC an animal. The compositions comprise an immunising amount of a  
CC recombinant B. anthracis protective antigen (rPA) protein and a  
CC recombinant B. anthracis lethal factor (rLF) protein. The composition and  
CC methods are useful in protecting against anthrax or lethal infections  
CC caused by B. anthracis. The present sequence is pBPI backbone vector DNA  
CC used in the invention.  
XX  
SQ Sequence 6694 BP; 2086 A; 1229 C; 1544 G; 1835 T; 0 U; 0 Other;

Query Match 27.5%; Score 867; DB 12; Length 6694;  
Best Local Similarity 100.0%; Pred. No. 2.3e-205;  
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1475 TGCTTAGGAACAGAGTTATTAATAGCTGAATGAAGACGGTGCTCTCCAAATATTTCTAT 1534  
Db 5219 TGCTTAGGAACAGAGTTATTAATAGCTGAATGAAGACGGTGCTCTCCAAATATTTCTAT 5278

QY 1535 TTAGAAAAGCAATCTAAAATTTCTGAAAAGGAATGAGAATAGTGAATGGACCAATAA 1594  
Db 5279 TTAGAAAAGCAATCTAAAATTTCTGAAAAGGAATGAGAATAGTGAATGGACCAATAA 5338

QY 1595 TAATGACTAGAGAAGAAGAAATGGAATTTGTCATGAAATTAAGGACGAATATTGGATA 1654  
Db 5339 TAATGACTAGAGAAGAAGAAATGGAATTTGTCATGAAATTAAGGACGAATATTGGATA 5398

QY 1655 AATATGGGGATGATGTTAAGGCTATTGGTGTATTTATGGCTCTCTTGGTCGTCAGACTGATG 1714  
Db 5399 AATATGGGGATGATGTTAAGGCTATTGGTGTATTTATGGCTCTCTTGGTCGTCAGACTGATG 5458

QY 1715 GGCCCTATTCCGATATTGAGATGATGTGTCTCATGTCAACAGAGGAAGCAGAGTTCCAGCC 1774  
Db 5459 GGCCCTATTCCGATATTGAGATGATGTGTCTCATGTCAACAGAGGAAGCAGAGTTCCAGCC 5518

QY 1775 ATGAATGGACAACCCGGTGAAGAGTGAAGTGAATTTTGTATAGCAAGAGATTCTTAC 1834  
Db 5519 ATGAATGGACAACCCGGTGAAGAGTGAAGTGAATTTTGTATAGCAAGAGATTCTTAC 5578

QY 1835 TAGATTATGCATCTCAGGTGGAATCAGATTGGCGGCTTACACATGGTCAATTTTCTCTA 1894  
Db 5579 TAGATTATGCATCTCAGGTGGAATCAGATTGGCGGCTTACACATGGTCAATTTTCTCTA 5638

QY 1895 TTTTGGCGAATTTATGATTCAGGTGGATACCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 1954  
Db 5639 TTTTGGCGAATTTATGATTCAGGTGGATACCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 5698

QY 1955 TAGAAGCCCAACGTTCCACGATGCCGATTTGTGCCCTTATCGTAGAAGAGCTGTTGAAT 2014  
Db 5699 TAGAAGCCCAACGTTCCACGATGCCGATTTGTGCCCTTATCGTAGAAGAGCTGTTGAAT 5758

QY 2015 ATGAGGCAAAATGGCGTAAATATTCTGTGCAAGGACCGACAACATTTCTACCATCCTTGA 2074  
Db 5759 ATGAGGCAAAATGGCGTAAATATTCTGTGCAAGGACCGACAACATTTCTACCATCCTTGA 5818

QY 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGGTCTGCGATCATCGCATCTGTTATA 2134  
Db 5819 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGGTCTGCGATCATCGCATCTGTTATA 5878

QY 2135 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTTATG 2194  
Db 5879 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTTATG 5938

QY 2195 ACCATCTGTGCCAGTTCTGTAATGTTGTTCAACTTTCCGACTCTGAGAACTTCTGGAAT 2254  
Db 5939 ACCATCTGTGCCAGTTCTGTAATGTTGTTCAACTTTCCGACTCTGAGAACTTCTGGAAT 5998

QY 2255 CGCTAGAGAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACACGCGATATATAGTGG 2314  
Db 5999 CGCTAGAGAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACACGCGATATATAGTGG 6058

QY 2315 ATGTGTCAAAACGCATACCACTTTTGAA 2341  
Db 6059 ATGTGTCAAAACGCATACCACTTTTGAA 6085

Search completed: February 11, 2006, 18:11:22  
Job time : 1228 secs

***THIS PAGE BLANK (uspto)***

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2006, 18:11:35 ; Search time 8017 Seconds

(without alignments)  
18418.365 Million cell updates/sec

Title: US-10-681-086-1

Perfect score: 3156

Sequence: 1 ggatccagcaggttacgagc.....atcggtatgatgcgaattc 3156

Scoring table: IDENTITY NUC

Gapop 10\_0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_est3:\*

4: gb\_est4:\*

5: gb\_est5:\*

6: gb\_est6:\*

7: gb\_est7:\*

8: gb\_est8:\*

9: gb\_est9:\*

10: gb\_est10:\*

11: gb\_est11:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	175.2	5.6	1064	10 AY421133	Mus muscu
2	175.2	5.6	1607	4 AK087433	Mus muscu
3	175.2	5.6	1677	4 AK075782	Mus muscu
4	173.6	5.5	1064	10 AY421131	Homo sapi
5	173.6	5.5	2327	4 AK013138	Mus muscu
6	171.4	5.4	1737	4 AK076469	Mus muscu
7	167.2	5.3	1528	4 CR860400	Pongo pyg
8	161.4	5.1	911	5 BX325788	BX325788
9	159.8	5.1	794	6 CA778095	ip17h09.x
10	159.8	5.1	810	6 CA778022	ip16h09.x
11	157.6	5.0	862	8 CX934609	JGI CAAR3
12	157.4	5.0	805	8 CX438675	JGI_XZG67
13	157.4	5.0	1064	10 AY421132	Pan trogl
14	157.2	5.0	787	2 BG752866	602732509
15	157.2	5.0	871	8 DN019683	JGI CAAR2
16	157.2	5.0	901	8 DR871972	JGI CABG1
17	156	4.9	897	8 DN022379	JGI CAAR3
18	155.8	4.9	844	8 DN080099	JGI_CABD1
19	155.6	4.9	885	8 DN083438	JGI_CABD1
20	155.6	4.9	906	8 CX389898	JGI_XZT37
21	155.4	4.9	609	7 CF900218	A0312H09-
22	155.4	4.9	609	7 CF900995	A0322E10-

23	155.4	4.9	615	7 CF900454	CF900454 A031SG03-
24	155	4.9	677	2 BG928003	BG928003 HNC45-1-H
25	154.4	4.9	1004	2 BB610357	BB610357 BB610357
26	154.2	4.9	895	2 BI251895	BI251895 602952509
27	154	4.9	978	5 BU519273	BU519273 AGENCOURT
28	153.8	4.9	792	7 CV675649	CV675649 i631807.k
29	153	4.8	797	2 BG747688	BG747688 602705159
30	152.6	4.8	580	7 CN347398	CN347398 170005830
31	152.2	4.8	789	8 DN027048	JGI CAAR6
32	152.2	4.8	864	8 DN028149	DN028149 JGI CAAR6
33	152.2	4.8	874	8 DN025254	JGI CAAR5
34	152.2	4.8	883	7 CR422328	CR422328 CR422328
35	152.2	4.8	891	7 CR567595	CR567595 CR567595
36	152	4.8	908	6 CB566539	CB566539 AGENCOURT
37	151.2	4.8	733	5 BX325548	BX325548 BX325548
38	151.2	4.8	867	8 DN025549	JGI CAAR5
39	150.8	4.8	955	5 BU526482	BU526482 AGENCOURT
40	150.6	4.8	781	8 CX327170	JGI_XZT15
41	150.6	4.8	786	8 CX969308	JGI_CAAP3
42	150.6	4.8	882	8 CX362459	CX362459 JGI_XZT42
43	150.6	4.8	884	8 CX468737	CX468737 JGI_XZG33
44	150.6	4.8	889	8 DN033808	JGI CAAR1
45	150.6	4.8	911	8 CX391320	CX391320 JGI_XZT38

ALIGNMENTS

RESULT 1	AY421133	1064 bp	DNA	linear	GSS 17-DEC-2003
LOCUS	Mus musculus GCAT gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.				
DEFINITION	AY421133.1 GI:39777090				
ACCESSION	GSS.				
VERSION	Mus musculus (house mouse)				
KEYWORDS	Mus musculus				
SOURCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.				
ORGANISM	1 (bases 1 to 1064)				
REFERENCE	Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.				
AUTHORS	Inferring nonneutral evolution from human-chimp-mouse orthologous				
TITLE	Gene trios				
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
PUBMED	14671302				
REFERENCE	2 (bases 1 to 1064)				
AUTHORS	Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA				
COMMENT	This sequence as made by sequencing genomic exons and ordering them based on alignment.				
FEATURES	Location/Qualifiers				
source	1..1064				
	/organism="Mus musculus"				
	/mol_type="genomic DNA"				
gene	/gb_xref="taxon:10090"				
	<1..>1064				
	/gene="GCAT"				
	/locus_tag="HCM7455"				
ORIGIN					

Query Match 5.6%; Score 175.2; DB 10; Length 1064;  
Best Local Similarity 49.9%; Pred. No. 7.5e-35;  
Matches 475; Conservative 0; Mismatches 468; Indels 9; Gaps 1;

Qy 158 CAAACAATTATTTAGGCTCCGAAGCATAGACGTTTGTATCGATCGACCCCAACACGAT 217  
Dy |||||  
19 CCAATAACTACTCGGCTGAGACGCCACCTCGAGTGATCCAGCGAGTCTGCAGACTC 78  
Qy 218 TGCAGCAATTTGGACAGGAAGACGCGTTCACGTTTAAACACACAGCAATTCGCTCTGGC 277  
Dy |||||  
79 TGGAGGATTTGGAGCTGACCTCAGTTCTCATCTCGATTATCTGTGGACCCAGAGCATCC 138  
Qy 278 ATGAAAGCTAGAAAAGAGATTCGCAGCTTTAACTGACAGAACGCGCCCTCTGTTTT 337  
Dy |||||  
139 ATAAGAATCTAGAAGCCCAAGATAGCCCACTTCCACGACGCTGAGGACGCCATCCTCTATC 198  
Qy 338 CGAGCGTTACTTGGCCCAATGCTGGTCTCTTTCATCTTCCGCAAGAAAGCAAGATGCA 397  
Dy |||||  
199 CCAGCTGTTTCGATGCCCAACCTGGCTATTATTAGGCTTTTGTCTACCCACGAGATCGG 258  
Qy 398 TTTTAAAGTGACCAAGCTCAATCATGCAAGATGATGACGCGCTGCCGACTTTCTAAGGCTG 457  
Dy |||||  
259 TTCTGTCAATGAGCTGNAACCATGCTTCCATCATTGACGGCATCCGACTGTGCAAGGCC 318  
Qy 458 ATACAGTTGTTTATCGGCATATTGATATGAATGATCTTTGAAAAACAAGCTGAATGAACAC 517  
Dy |||||  
319 ACAAGTACCGTTACCGCCACCTCGACATGGCTGACCTGGAAGCCAAAGCTGAAGGAGGCTC 378  
Qy 518 AGCGTTATCAGCGCGCTTTTATCTGTAACAGACGAGTATTCAGCATGATGGCAACATCG 577  
Dy |||||  
379 AGAAGCAGAGCTCGCCCTGGTGGCCACAGATGAGCTTTCTCCATGATGGTGACATCG 438  
Qy 578 CCCCTCTTGATCAGATCATCTCACTTGCAGAACGCTATCATGCGCTTGTGTCGTGTGATG 637  
Dy |||||  
439 CTCCCTACAGGACATCTGCGCCTTTCGCGCCCAATATGGTGGCTGGTCTTGTGGATG 498  
Qy 638 ATGCCACGCAACAGGAGTTTGGCGGATTCGGGACAAAGAAACGAGTGAATCTTTGGTG 697  
Dy |||||  
499 AATGCCATGCCATGGCTTCTTGGGCCCCACAGACGCGGCACAGATGAGCTGTAGTGTG 558  
Qy 698 TTTGTCCGACATGTTATCGGCACCTTAAGCAAG-----CTGTTGGCGCGGAAG 748  
Dy |||||  
559 TCATGGACCAAGTCAACATATCACTCACTCCCTCGGGGAAGCGCTGGTGGAGCATCAG 618  
Qy 749 GAGGTTTTTGGCGGAGGATCAGCGTCTTCATCGACTTTTTCCTGAAACCATGCCAGAACAT 808  
Dy |||||  
619 GGGATATACACCGGCGCTGAGCACTGGTGTCCCTGCTACGCGCAGCGTTCTCGGCCCT 678  
Qy 809 TTATCTTTCAAACCGCTATTTCGGCAGCGACTGTGGGCTGCTCAGAGGCTTTCAACA 868  
Dy |||||  
679 ACCTCTTCTCAACAGCGCTGCCCTGCTGTAGTCGGCTGTCTTCTAAGGCGCCTGGACC 738  
Qy 869 TCATTGAAGCAGCAGGGAACAGCAGACTTTTATTTCTTATATCAGCATGATCAGAA 928  
Dy |||||  
739 TGCTCATGGAGAGCAATGCCATCATCCAGTCTATGGCTGCCAAGACCGCGGTTCCGCA 798  
Qy 929 CCAGTCTGAAGATATGGGTTATGTGTGAAAGAGATCACAACCGGATTTATCTCTGTAG 988  
Dy |||||  
799 GTAAGATGGAAGCTGTGTTTACTGTCTCAGAGCGGATCACCCCATCTGCGCCTGTGA 858  
Qy 989 TCATTGGCGATGCCCAATAAAGCGTCTTATTTGCTGAAAAACTGACGGCAAGGGAATTT 1048  
Dy |||||  
859 TGCTGGGAGATGCACGGTGTCTTCTCAATGGCAGATGACATGTTAAAGAAAGGTATCT 918  
Qy 1049 ATGCTCTGCCATTCGGCGCCCAACCGTTGCGCGGGTGAAGCGGATTCG 1100  
Dy |||||  
919 TTGTATCGGATTTAGCTACCTCTGTGGTCCCTAAGGGCAAGCGCCCGGATCCG 970

## RESULT 2

AK087433  
LOCUS 1607 bp mRNA linear HTC 03-APR-2004  
DEFINITION Mus musculus 0 day neonate eyeball cDNA, RIKEN full-length enriched library, clone:R130112P7 product:glycine C-acetyltransferase (2-amino-3-ketobutyrate-coenzyme A ligase), full insert sequence.  
ACCESSION AK087433

VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AK087433.1 GI:26352483  
HTC; CAP trapper.  
Mus musculus (house mouse)

## ORGANISM

Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

1 Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

10349636

## REFERENCE

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Sugahara, K.,

Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

11042159

## REFERENCE

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,

Konno, H., Akiyama, J., Nishi, K., Kiteunai, T., Tashiro, H., Itoh, M.,

Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

Yamanoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,

Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J.,

Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

11076861

## REFERENCE

4 The RIKEN Genome Exploration Research Group Phase II Team and the

FANTOM Consortium.

Functional annotation of a full-length mouse cDNA collection

Nature 409, 685-690 (2001)

## REFERENCE

5 The FANTOM Consortium and the RIKEN Genome Exploration Research

Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation

of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 1607)

## REFERENCE

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A.,

Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,

Furuno, M., Handgaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,

Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,

Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,

Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,

Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,

Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D.,

Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Sogabe, Y.,

Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,

Tejima, Y., Toya, T., Yamamura, T., Yamanaka, I., Yasunishi, A.,

Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of

Physical and Chemical Research (RIKEN), Laboratory for Genome

Exploration Research Group, RIKEN Genomic Sciences Center (GSC),

RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,

Kanagawa 230-0045, Japan (E-mail: genome-res@gsr.riken.jp,

URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,

Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

URL: http://fantom.gsc.riken.jp/

Location/Qualifiers

1. .1607

FEATURES  
source

November 2005

Published Applications Nucleic Acid and Published Applications Amino Acid database searches now generate two sets of results each. The Published Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published Applications New databases; older published applications make up the Published Applications Main databases.

Searches run against Nucleic Acid Published Applications produce two sets of results, with the extensions **.npbm** (Published Applications NA\_Main) and **.rnpbm** (Published Applications NA\_New).

Searches run against Amino Acid Published Applications produce two sets of results, with the extensions **.rapbm** (Published Applications AA\_Main) and **.rapbm** (Published Applications AA\_New).

***This Page Blank (uspto)***



		/organism="Mus musculus"			
		/mol_type="mRNA"			
		/strain="C57BL/6J"			
		/db_xref="FANTOM DB:EI30112F17"			
		/db_xref="taxon:10090"			
		/clone="EI30112F17"			
		/tissue_type="eyeball"			
		/clone_lib="RIKEN full-length enriched mouse cDNA library"			
		/dev_stage="0 day neonate"			
CDS		20..1270			
		/notes="unlabeled protein product; glycine			
		C-acetyltransferase (2-amino-3-ketobutyrate-coenzyme A			
		ligase) (MGI:1349389)			
		putative"			
		/codon_start=1			
		/protein_id="BAC39872.1"			
		/db_xref="GI:26352484"			
		/translation="MASFWHGALSPGRHSAALQRLCILDSLEGIRGATWSE			
		RVITSRQGSIRVDGSGGILNFCANNYLGLSSHPAVIQGLQTLSEFGAGLSSTRFI			
		CGTQSIRKLEAKIAHQREDAIYPSCFDNAGLFEALLPEDAVLSDLNHASII			
		DGIRLCAHKYRRLHDMADLEAKLEAKQRLRLVATDCAFSGMDGIIAPLDICRLA			
		AOYGLVFDDECHATFLGPTGRGTDELLGMDQVTIINSTLKGALGASGGVTTGPE			
		PLVSLARSRPVLFSNLSLPPAVVGCASKALDLMESNALIOSMAAKTRFRSKMEA			
		GFTVSGADRPICVPMGLDRLSSQADMDMLKGI FVIGFSYVVPVPGKARIRVQISAV			
		HSEBIDRRCVEAFVEVGRHLGALP"			
ORIGIN					
		Query Match 5.6%; Score 175.2; DB 4; Length 1607;			
		Best Local Similarity 49.9%; Pred. No. 8.4e-35;			
		Matches 475; Conservative 0; Mismatches 468; Indels 9; Gaps 1;			
QY	158	CAAAACAATATTATTTAGGCGTCGCAAGCGATGACAGCTTTTGATCGATCGAGCGCCAAACAGCAT	217		
DB	225	CCAATAACTACCTGGGCGCTGAGCAGCACCCCTCGAGTGATCCAGGCGAGTCTCGAGACTC	284		
QY	218	TGACGCAATTTGGCAGGAGCAGCGGTTACGTTTAAACGACGAGCAATTCGGTCTGGC	277		
DB	285	TGGAGGAGTTTGGAGCTGGACTCAGTTCTACTCGATTCTCTGGGACCCAGAGCATCC	344		
QY	278	ATCAAAAGCTAGAAAAGAGATTGCCAGCTTTAAACTGCACAGAGCGGCCCTCGTGT	337		
DB	345	ATAAGAACTCTAGAAAGCCAGATAGCCCACTTCCACGAGCGTAGGAGCCCATCTCTATC	404		
QY	338	CGAGCGGTACTTTGGCCCAATGTGCGGTCTCTTTTCATCTCTTCCAGAAAAGAGATGTCA	397		
DB	405	CCAGCTGTTTCGATGCCAACGCTGGCCCTATTCTGAGGCTTTTGCTGACCCAGAGAGTCGG	464		
QY	398	TTTAAAGTGACAGCTCAATCATCTGCAAGTATGATTCGACGCGCTCGGACTTTCTTAAGCTG	457		
DB	465	TTCTGTGAGATGAGCTGAACCATGCTTCCATCATTTGACGGCATCCGCACTGTGCAAGGCC	524		
QY	458	ATACAGTTGTTTATCGCATATTGATATGAATGATCTTTGAAAACAAGCTGAATGAACAAC	517		
DB	525	ACAAGTACCGTTACCGCCACTCGACATGCGTCACTTGGAAAGCCAGCTGAAGAGGCTC	584		
QY	518	AGCGTTATCAGCGCGGTTTTATCGTAAACAGACGAGGATTTTCAGCATGATGGCACAATCG	577		
DB	585	AGAAGCACAGCTCGCGCTGGTGCCACACAGATGAGCTTTCTCCATGATGGTGACATCG	644		
QY	578	CCCTCTTGATCAGATCATCTCATCTTGGCGAAAGCGTATCATGCGCTTCGGTGGTGTGATG	637		
DB	645	CTCCCTTACAGGACATCTGCGCGCTTGGCCGCCCAATATGTTGGTGGCTGTCTTTGTGATG	704		
QY	638	ATGCCACGCAACAGGAGTTTGGCGATTTCGGGCAAGCAAGCAAGTGAATCTTTGGTG	697		
DB	705	AATGCCATGCTCTGGCTTTCTTGGGCCCAAGACGCGGCAAGATGAGCTGCTAGGTG	764		
QY	698	TTTGTCCCGACATTGTTATCGGCACCTTAAGCAAG-----CTGTTGGCGCGGAAG	748		
DB	765	TCATGACACAGGTACCATCATCACTCCACCTGGGAGGCGCTGGTGGAGCATCAG	824		
QY	749	GAGGTTTGGCGGAGGATCAGCGTCTTCATTCGACTTTTTCGTGAACCATGCGCAGAACAT	808		
Db	825	GGGGATATACACCGGGCGCTGAGCCACTGGTGTCCCTGTCTACGGCAGCGTTCTCGGGCCT	884		
QY	809	TTATCTTTCAAACCGCTATTTCGCCAGCCAGCTGTGGCTGTCTACGAGGCTTTCAACA	868		
Db	885	ACCTCTTCTCCACACGCTGCCCTGTCTGTAGTCGGCTGTGCTTCTAAGCCCTGGACC	944		
QY	869	TCATTGAAGCAGCAGGAAAAACGACAGCTTTTATTTCTTATATCAGCATGATCAGAA	928		
Db	945	TGCTCATGAGAGCAATGCCATCATCCAGTCTATGGTCTGCAAGAGCCCGCGGTTCGCA	1004		
QY	929	CCAGTCTGAAGAATATGGGTATGTGTGAAGAGATCACACCGATTATTCCTGTAG	988		
Db	1005	GTAAGATGAAGCTGCTGGTTTTTCTGTCTCAGAGCCGATCACCCCATTCGCCCTGTGA	1064		
QY	989	TCATTGGCGATGCCATAAAACGGTCTCTATTTCCTGAAAACTGCAGGCGCAAGGAATTT	1048		
Db	1065	TGCTGGAGATGCACGGTTGTCTTCTCAATGCGCAGATGACATGTTAAAGAAAGTATCT	1124		
QY	1049	ATGCTCTGCGCATTCGGCGCCCAACCGTTGCGCGCGGTGAAAGCCGGATTGCG	1100		
Db	1125	TTGTCTATCGAATTTAGCTACCTGTGCTCCCTTAAGGGCAAGGCCCGGATCCG	1176		
RESULT 3		1677 bp mRNA linear HTC 03-APR-2004			
AK075782		Mus musculus 10 day old male pancreas cDNA, RIKEN full-length			
LOCUS		enriched library, clone:1810021L23 product:glycine			
DEFINITION		C-acetyltransferase (2-amino-3-ketobutyrate-coenzyme A ligase),			
		full insert sequence.			
ACCESSION		AK075782			
VERSION		AK075782.1 GI:26096539			
KEYWORDS		HTC; CAP trapper.			
SOURCE		Mus musculus (house mouse)			
ORGANISM		Mus musculus			
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
		Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;			
		Sciurognathi; Muridae; Murinae; Mus.			
REFERENCE		1			
AUTHORS		Carninci, P. and Hayashizaki, Y.			
TITLE		High-efficiency full-length cDNA cloning			
JOURNAL		Meth. Enzymol. 303, 19-44 (1999)			
PUBMED		10349636			
REFERENCE		2			
AUTHORS		Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,			
		Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.			
TITLE		Normalization and subtraction of cap-trapper-selected cDNAs to			
JOURNAL		Prepare full-length cDNA libraries for rapid discovery of new genes			
PUBMED		Genome Res. 10 (10), 1617-1630 (2000)			
REFERENCE		3			
AUTHORS		Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,			
		Konno, H., Akiyama, J., Nishi, K., Kitsuunai, T., Tashiro, H., Itoh, M.,			
		Sumi, N., Ihii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,			
		Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,			
		Fujiiwake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,			
		Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,			
		Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.			
TITLE		RIKEN integrated sequence analysis (RISA) system--384-format			
JOURNAL		sequencing pipeline with 384 multicapillary sequencer			
PUBMED		Genome Res. 10 (11), 1757-1771 (2000)			
REFERENCE		4			
AUTHORS		The RIKEN Genome Exploration Research Group Phase II Team and the			
		FANTOM Consortium			
TITLE		Functional annotation of a full-length mouse cDNA collection			
JOURNAL		Nature 409, 685-690 (2001)			
REFERENCE		5			
AUTHORS		The FANTOM Consortium and the RIKEN Genome Exploration Research			
		Group Phase I & II Team.			
TITLE		Analysis of the mouse transcriptome based on functional annotation			
JOURNAL		of 60,770 full-length cDNAs			
		Nature 420, 563-573 (2002)			

```

REFERENCE
AUTHORS
6 (bases 1 to 1677)
Adachi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
Arakawa,T., Bono,H., Carninci,P., Fukuda,S., Fukunishi,Y.,
Furuno,M., Hanagaki,T., Hara,A., Hayatsu,N., Hiramoto,K.,
Hiraoka,T., Hori,F., Inotani,K., Ishii,Y., Itoh,M., Izawa,M.,
Kasukawa,T., Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K.,
Nomura,K., Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C.,
Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
Shibata,K., Shibata,Y., Shingawa,A., Shiraki,T., Sogabe,Y.,
Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Tanaka,T.,
Tejima,Y., Toya,T., Yamamura,T., Yamanaka,I., Yasunishi,A.,
Yoshida,K., Yoshino,M., Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan [E-mail: genome-res@ac.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216]
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.
Location/Qualifiers
1. 1677
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:1810021L23"
/db_xref="taxon:10090"
/clone="1810021L23"
/sex="male"
/tissue_type="pancreas"
/closure_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="10 day old"
5. 1251
/notes="glycine C-acetyltransferase
(2-amino-3-ketobutyrate-coenzyme A ligase)
(MGI:1349389)
putative"
misc_feature
ORIGIN
Query Match 5.6%; Score 175.2; DB 4; Length 1677;
Best Local Similarity 49.9%; Pred. No. 8.5e-35;
Matches 475; Conservative 0; Mismatches 468; Indels 9; Gaps 1;
QY 158 CAACAATTATTTAGGCTCCAGCGATAGACGTTTGATGATGATGAGCCCAACAGCAT 217
Db |||||
QY 206 CCAATAACTACCTGGCTGAGCAGCACCCTCGAGTGATCCAGGAGGCTGCGAGACTC 265
Db |||||
QY 218 TGCAGCAATTTGGACAGGAGCAGCGGTTCACGTTTAAACAGCAGGCAATTCGTCTGGC 277
Db |||||
QY 266 TGGAGGAGTTTGGAGCTGGACTCAGTTCTACTCGATTATTCGTGGGACCCAGAGCATCC 325
QY 278 ATGAAAAGCTAGAAAAGAAATGCCAGCTTTAACTGACAGAGCGCCCTGCTGTTTT 337
Db |||||
QY 326 ATAGATCTAGAGCCAGNATAGCCACCTTCACAGCGTGAGAGCCCATCTCTATC 385
QY 338 CGACGGTTACTTGGCCAATGTCGGTGTCTTTTCATCTCTGCCAGAAAGGAAGATGCA 397
Db |||||
QY 386 CCAGCTGTTTCGATGCCAACGCTGGCTGCTATTTAGGCTTTGCTGACCCAGAGATCGG 445
QY 398 TTTTAAGTGACCACTCAATCATCATGATGATGTCGAGCGCTGCGGACTTCTTAAGCTG 457
Db |||||
QY 446 TTCTGTGAGATGATGAACCATCTTCCATCATTTGACGGCATCCGACTGTGCAAGCCC 505
QY 458 ATACAGTTGTTTATCGGCATATTGATGATGATGATCTTTGAAACCAAGCTGAATGAACAC 517

```

```

Db |||||
506 ACAGTACCGCTTACCGCCACTGGACATGCTGACCTGACCTGGAAGCAAGCTGAAGAGGCTC 565
QY |||||
518 AGCGTTATCAGCGCGGTTTATTCGTAAACAGACGAGGATTTTCAGCATGATGATGCAATCG 577
Db |||||
566 AGAAGCACAGCGCTCGCGCTGCTGGCCACAGATGAGAGCTTTCTCCATGATGATGATGATCG 625
QY |||||
578 CCCTCTTGTATCAGATCATCTCACTTTCGGAACGCTATCATGCTTCTGCTGCTGCTGATG 637
Db |||||
626 CTCCCTTACAGGACATCTCGCGCTTCCCGCCCAATATGGTGCCTGCTTTTGTGGATG 685
QY |||||
638 ATGCCCAACGCAACGAGGATTTTGGCGGATTTTCGGGACCAAGAAACGAGTGAATCTTTGGTG 697
Db |||||
686 AATGCCATGCACTGGCTTTCTTGGGCCCACAGAGCGGGGCACAGATGAGCTGCTAGGTG 745
QY |||||
698 TTTGTCCCGACATTTGTTATCGGCACCTTAAGCAAG-----CTGTTGGCGCGGAAG 748
Db |||||
746 TCATGGACACAGGTCACCATCATCAACTCCACCCTGGGGAAGGCGCTGGGTGGAGCATCAG 805
QY |||||
749 GAGGTTTGGCGCAGGATCAGCGGTCTTCATCGACTTTTTCGTAACCATGCCAGAACAT 808
Db |||||
806 GGGGATATACACCGGCGCTGAGCGCACTGGTGTCTCTGCTACGGCAGCGTTCTCGGCCCT 865
QY |||||
809 TTATCTTTCAAAACCGCTATTTCGCGCCAGCAGCTGTGCGGCTGCTCAGAGGCTTTCAACA 868
Db |||||
866 ACCTCTTCTCAACAGCGCTGCGCCCTGCTAGTGGCTGCTTCTTAAGGCCCTGGACC 925
QY |||||
869 TCATTGAAGCCAGCAGGGAACAGCAGCTTTTATTTTCTTATATCAGCATGATCAGAA 928
Db |||||
926 TGCTCATGAGAGCAATGCCATCATCCAGTCTATGGCTGCGCAAGACCGCGGTTCGCA 985
QY |||||
929 CCAGTCTGAAGATATGGGTTATGTGTGAAGAGAGATCACACCCGATATTCTCTGTAG 988
Db |||||
986 GTAAGATGGAAGCTGTGGTTTACTGTCTCAGAGCGCGATCACCCCATCTGCCCTGTGA 1045
QY |||||
989 TCATTGGCGATGCCCATATAAACCGTCTTATTTGCTGAAAAAATCAGGCGCAAGGGAATTT 1048
Db |||||
1046 TGCTGGAGATGACGCGTTCTTCTCAATGCGAGATGACATGTTAAAGAAAGTATCT 1105
QY |||||
1049 ATGCTCTGCGCATTCGCGCCGCCAACCGTTGCGCGGTGTAAGCCGATTCG 1100
Db |||||
1106 TTGTCATCGGATTTAGTACCTCTGTGTCCTTAAGGCAAGCGCCCGATCCG 1157

RESULT 4
AY421131 1064 bp DNA linear GSS 17-DEC-2003
LOCUS Homo sapiens GCAT gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY421131
VERSION AY421131.1 GI:39777088
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominidae; Homo.
1 (bases 1 to 1064)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Givello,D.R., Lu,F., Murphy,B.,
Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 1064)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Givello,D.R., Lu,F., Murphy,B.,
Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,

```



JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/, Tel:81-45-503-9222, Fax:81-45-503-9216)	
COMMENT	Please visit our web site (http://genome.gsc.riken.jp/) for further details. cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5', GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 7.5 and subtraction to Rot = 37.5. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGATTCGAGTTCAATTAATCCCTCCCTCCCTCC 3'], cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI. Host: SOUR.	
FEATURES	Location/Qualifiers	
source	1..2327 /organism="Mus musculus" /mol_type="mRNA" /strain="C57BL/6J" /db_xref="FANTOM DB:2810422022" /db_xref="taxon:10090" /clone="2810422022" /tissue_type="whole body" /clone_lib="RIKEN full-length enriched mouse cDNA library" /dev_stage="10, 11 days embryo" 16..1266 /note="unnamed protein product; glycine C-acetyltransferase (2-amino-3-ketobutyrate-coenzyme A ligase) (MGD MG1:1349389) putative" /codon_start=1 /protein_id="BAB28673.1" /db_xref="GI:12850315" /translation="MWASFMWHGALSFGRRASALQRLCILDSELEGIRGAGTWKSE RVITSROGSPRVDSIGILNFCANNVGLSSHPAVIOAGLOTLEFFGAGLSSTRI CGTOSILHKEAKIAHFHOREDAIILYPCFDANAGLFEALLTPEDAVLSDELNHASII DGLRCKAHKRYRHLMDADLEAKLEAKRHLRLVATDGFSDMGDIAPLODLCRLA AQYALVFDSECHATGFLGPTGRGDELLGVMDQVTIINSTGLKAGGAGGHTTGE PLVLSLRQSRPYLFSNLSPPAVVGCASKALDILMESNAIIQSMAAKTRFRSRXMEAA GFTVSGADHPICPVMLGDARLSSQWADMLKKGIFVIGFSYPVVPVPGKARIRVQISAV HSEEDIDRCVEAFVEVGRHLGALP"	
CDS	/note="unnamed protein product; glycine C-acetyltransferase (2-amino-3-ketobutyrate-coenzyme A ligase) (MGD MG1:1349389) putative" /protein_id="BAB28673.1" /db_xref="GI:12850315" /translation="MWASFMWHGALSFGRRASALQRLCILDSELEGIRGAGTWKSE RVITSROGSPRVDSIGILNFCANNVGLSSHPAVIOAGLOTLEFFGAGLSSTRI CGTOSILHKEAKIAHFHOREDAIILYPCFDANAGLFEALLTPEDAVLSDELNHASII DGLRCKAHKRYRHLMDADLEAKLEAKRHLRLVATDGFSDMGDIAPLODLCRLA AQYALVFDSECHATGFLGPTGRGDELLGVMDQVTIINSTGLKAGGAGGHTTGE PLVLSLRQSRPYLFSNLSPPAVVGCASKALDILMESNAIIQSMAAKTRFRSRXMEAA GFTVSGADHPICPVMLGDARLSSQWADMLKKGIFVIGFSYPVVPVPGKARIRVQISAV HSEEDIDRCVEAFVEVGRHLGALP"	
ORIGIN	RESULT 6 AK076469 LOCUS DEFINITION Mus musculus 0 day neonate head cDNA, RIKEN full-length enriched library, clone:483341S122 product:glycine C-acetyltransferase (2-amino-3-ketobutyrate-coenzyme A ligase), full insert sequence. AK076469 ACCESSION VERSION AK076469.1 GI:26096824 KEYWORDS HTC; CAP trapper. SOURCE Mus musculus (house mouse) ORGANISM Mus musculus REFERENCE 1 Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) 10349636 REFERENCE 2 Carninci, P., Shibata, Y., Hayateu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)	
Qy	158	CRAACAATTATTAGGCTCGCAGCAGTAGAGTTCGATCGATCGACGCCAACACCAT 217
Db	221	CCAATAACTACTGGGCTTGAGCGCCACCTGCGAGTATCCAGCGAGGTCGAGACTC 280
Qy	218	TGCAGCAATTGGGACAGGACGAGCGGTTACCGTTAAACGACAGGCAATTCGGTCTGGC 277
Db	281	TGGAGGAGTTTGGAGCTGGACTCAGTTCTACTCGATTATCTGTGGGACCCAGAGCATCC 340
Qy	278	ATGAAAAGCTAGAAAAGAGATTGCCAGCTTTAACTGACAGACGCGCCCTGCTGTTTT 337
Db	341	ATAAGAATCTAGAAGCCAGATAGCCCACTTCCACGAGCGTGGAGCGCCATCCTCTATC 400
Qy	338	CGAGCGGTTACTTGGCCATGTCGCTGTCTTTCATCTCCGAGAAAGAGAGATGCA 397
Db	401	CCAGCTGTTTCGATGCCAACGCTGCGCTTATTGAGGCTTTGCTGACCCAGGAGATCGG 460
Qy	398	TTTTAAGTGACCACTCAATCATGCAAGTATGATCGAGCGGTGCGGACTTCTTAAGGCTG 457
Db	461	TTCTGTGATGAGCTGAACCATGCTTCCATCATTTGACGGCATCCGACTGTGCAAGGCC 520
Qy	458	ATACAGTTGTTTATCGGCATATTGATATGATATGATATGATATGATATGATATGATATG 517
Db	521	ACAAGTACCGTTACCGCCACCTGGACATGGCTGACCTGGAAAGCCAAAGCTGAAGGAGGCTC 580
Qy	518	AGCGTTATCAGCGCGGTTTTATCGTAACAGACGAGGAGTATTACAGCATGGATGGACAAATCG 577
Db	581	AGAAGCACAGGCTCGCGCTGGTGCCACAGATGAGGCTTTCTCCATGGATGGTGCATCG 640
Qy	578	CCCTCTTTGATCAGATCATCTCACTTGGGAAAACGCTATCATGCTTCTGTTGCTGTTGATG 637
Db	641	CTCCCTTACAGGACATCTGCGCGCTTCCGCGCCAAATATGTTGCGCTTCTTGTGTGGATG 700
Qy	638	ATGCCACGCAACAGGAGTTTGGGCGATTTCGGGCAATTCGGGACAAAGAAACGAGTGAATATCTTTGGTG 697
Db	701	AATGCCATGCCACTGGGCTTCTTGGGCGCCACAGACGCGGCGACAGATGAGCTGCTAGGTG 760
Qy	698	TTTGTCCGACATTGTTATCGGCACCTTAAAGCAAG-----CTGTTGGCGCGGAAG 748
Db	761	TCATGGACAGGTCAACATCATCACTCCACCTCGGGGAAGGCGCTGGTGGAGCATCAG 820
Qy	749	GAGGTTTTGCGGAGGATCAGCGTCTTTCATCGACTTTTTCGCTGAACCATGCCAGAACAT 808
Db	821	GGGACATACACCGCGCTGAGCCACTGGTGCTCCTGCTACGCGACGCTTCTCGGCGCT 880
Qy	809	TTATCTTTTCAACCGCTATTTCGCCACGACGCTGTCGGCTGCTCAGAGGCTTTTCAACA 868
Db	881	ACCTCTTCTCCAAAGCGCTGCCCTGCTGTAGTCGCTTCTTAAGGCGCTTGA 940
Qy	869	TCATTGAAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 928
Db	941	TGCTCATGGAGAGCAATGCCATCATCACTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1000
Qy	929	CCAGTCTGAAAGATATGGTTATGTGTGAAGAGATACACACCGATTTATCTCTGTAG 988
Db	1001	GTAAGATGGAAGCTGCTGGTTTTACTGCTCAGGAGCGGATCAACCCCATCTGCCCTGTGTA 1060
Qy	989	TCATTGGCGATGCCATAAAGCGTCTTATTTGCTGAAAAAAGTGCAGGGCAAGGGAATTT 1048
Db	1061	TGCTGGGAGATGACACGGTGTGTTCTTCAATGGCAGATGACATGTTAAAGAAAGGTATCT 1120
Qy	1049	ATGCTCTCGCATTCGGCGCGCAACCGTTGCGCGGTTGAAGCCGATTCG 1100
Db	1121	TTGTCATCGGATTTAGCTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1172



Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Pongo.  
1 (bases 1 to 1528)  
Pouetka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,  
Mewes,H.W., Weill,B., Amlid,C., Osanger,A., Fobo,G., Han,M. and  
Wiemann,S.  
The German cDNA Consortium  
Direct Submission  
Submitted (12-NOV-2004) MIPS, Ingolstaedter Landstr.1, D-85764  
Neuerberg, GERMANY  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
Heidelberg/Germany) within the cDNA sequencing consortium of the  
German Genome Project. This clone (DKFZp459D1523) is available at  
the RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH in  
Berlin, Germany. Please contact RZPD for ordering:  
http://www.rzpd.de/cgi-bin/products/cl.cgi?CloneID=DKFZp459D1523  
Further information about the clone and the sequencing project is  
available at http://mips.gsf.de/projects/cdna/.  
Location/Qualifiers  
1..1528  
/organism="Pongo pygmaeus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9600"  
/clone="DKFZp459D1523"  
/tissue\_type="cortex"  
/clone\_lib="459 (synonym: pcor1). Vector pSport1\_Sfi; host  
DH10B; sites SfilA + SfilB"  
/dev\_stages="adult"  
/note="2-amino-3-ketobutyrate coenzyme A ligase,  
mitochondrial precursor (Homo sapiens)"  
1..1528  
/gene="DKFZp459D1523"  
24..1283  
/gene="DKFZp459D1523"  
/codon\_start=1  
/product="hypothetical protein"  
/protein\_id="CAH92526.1"  
/db\_xref="GI:55731640"  
/translation="MWAGNRAAVFVWSRRRAQSAALQRLGLEBELEGIRAGTW  
KSERVTSRQGFPHICVDGAVGGIINFANNYLGLSSHPVEVIOAGLOALEEFGALSSV  
RFICGTOSIHKNLAKIARPHOREDAIILPSCFDANAGLPEALITPDVAIISDELNHA  
SIIDGIRLCKAHKYRHLMDLEAKLOBAKHRLRLVATDGAFFSMDGDIAPLOEIC  
PLASRYGALVFDSCATGFLGPTGRGTDLLGYMDQVITLINSITLGAAGAGSGYTT  
CPGPIVSLLRARAPYLFNSLPPAVVGCASKALDLMESNTIIVQSMMAKTFQFRSRM  
EAGFTTISGASHPICPVNLGDARPASRWADMLKRGIFVIGFSYFVVPVPGKARIRVQI  
SAVHSEEDIDCCVEAFVEVGRHLGALP"  
ORIGIN  
Query Match 5.3%; Score 167.2; DB 4; Length 1528;  
Best Local Similarity 49.4%; Pred. No. 1.1e-32;  
Matches 470; Conservative 0; Mismatches 473; Indels 9; Gaps 1;  
Qy 158 CAAACAATATTATTAGGCTCGCAAGCATAGAGTTTCATCGATCGACCCCAACAGCAT 217  
Db 238 CCAACAACTACTGGCCCTGAGCACCCTGAGTGATTCAGCAGGCTCGCAGGCTC 297  
Qy 218 TGCAGCAATTTGGGACAGGACGCGGTTCACGTTTAAACGACAGGCAATTCGCTCGC 277  
Db 298 TGGAGGAGTTGGAGCTGGCCTCAGCTCTGCTCGGCTTTATCTGTGGAACCCAGAGCATCC 357  
Qy 278 ATGAAGACTAGAAAGAGATTCGACGTTTAACTGACAGAGCGCCCTGCTGTTT 337  
Db 358 ACAAGAATCTAGAAAGCAAAAATAGCCGCTTCCACAGCGGAGGATGCCATCTCTATC 417  
Qy 338 CGAGCGGTTACTTGGCCAAATGTCGGTGTCTTTTCATCTTCCAGAAAGGAAGATGTCA 397  
Db 418 CCAGCTGTTTTCAGCCCAACCGCGCTCTTTGAGGCCCTCTGTCNCCCGAGGACGAG 477  
Qy 398 TTTTAAGTGACAGCTCAATCATGCAAGTATGATCGAGCGGTGCGGACTTTCTTAAGGCTG 457  
Db 478 TCCTGTGACAGAGCTGAACCATGCTCCATCATCGAGGCATCCGGCTGTGCAAGGCC 537

Qy 458 ATACAGTTGTTTATCGGCATATTGATATGAATGATCTTGAAAAACAAGCTGAATGAACAC 517  
Db 538 ACAAGTACCGCTATCGCCACCTGGACATGGCTACCTAGAACCCAGCTGAGAGGCC 597  
Qy 518 AGCGTTATCAGCGCGCTTTTATTCGTAACAGACGAGGATTTTCAGCATGATGATGCAATCG 577  
Db 598 AGAAGCATCGCTGGCTAGTGGCCACTGATGGGCTTTTCCATGATGATGCGACATCG 657  
Qy 578 CCCCTTTGATCAGATCATCTCACTTGGGAACCTTATCATGCTTCTGTTGTTGATG 637  
Db 658 CACCCCTGCAAGAGATCTGCGGCTCTAGATATGCGGCTTTTCCATGATGATGCGACATCG 717  
Qy 638 ATGCCCAAGCAAGGAGTTTGGGCGATTTCGGGACAAAGAACGAGTGAATCTTTGTTG 697  
Db 718 AATCCCATGCCACCGGCTTCTGGGGCCACAGACGGGGCACAGACGAGCTGCTGGG 777  
Qy 698 TTTGTCCTCCGACATTTTATCGGCACCTTAAAGCAAG-----CTGTTGGCGCGAAG 748  
Db 778 TGATGGACCAAGTCAACCATCATCACTCCACCTGGGGAAGGCCCTGGGTGGAGCATCAG 837  
Qy 749 GAGGTTTGGCGGAGGATCAGCGTCTTCATCGACTTTTTCGTCGAACCATGCCAGAACAT 808  
Db 838 GGGGCTACAGACAGGCGCTGGGCGCTCTGGTGTCTCTGTCGCGACGCGCCCGGCCCT 897  
Qy 809 TTATCTTTCAAAACCGCTATTCCGCGCAGCAGCTGTGGGCTGTCTCACGAGCTTTTCAACA 868  
Db 898 ACCTCTTCTCAACAGTCTGCCACCTGCTGTTGGTGGCGCTCCNAGGCCCTAGATC 957  
Qy 869 TCATTGAAGCAGCAGGAGGAAAAACAGCAGCTTTTATTTTCTTATATCAGCATGATCAGAA 928  
Db 958 TGCTCATGGAGAGTAACACCATTTGTCAGTCTATGCTGCCAACAGACCCAGAGGTTCCGTA 1017  
Qy 929 CCAGCTCTGAAGATATCGGTTATGTTGTTGAAGAGATCACACCGGATTTATCTCTGTAG 988  
Db 1018 GTAAGATGGAAGCTGTCTGGCTTCACTATCTCGGAGCGCAGTCACCCCATCTGCCCCTGTGA 1077  
Qy 989 TCATTGGCGATGCCATAAAACCGTCTATTGCTGAAAAAATCTCAGGCGCAAGGAAATTT 1048  
Db 1078 TGCTGGGTGATGCCCGCGCGCTCTCGCATGGCGGATGACATGCTGNAAGAGGCATCT 1137  
Qy 1049 ATGCTCTGCCATTCGCGCCGCCAACCGTTGCGCGGGTGAAGCGCGGATTCG 1100  
Db 1138 TTGTCATCGGTTTCAGCTACCCCGTGTCCCAAGGCGCAAGGCCCGGATCCG 1189  
RESULT 8  
BX325788/c  
LOCUS  
DEFINITION BX325788 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens  
cDNA clone CS0DB006YD07 5-PRIME, mRNA sequence.  
ACCESSION BX325788  
VERSION BX325788.1 GI:30340511  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 911)  
Li,W.B., Gruber,C., Jesses,J. and Polayes,D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
7364.f





```
Qy 278 ATGAAAGCTAGAAAGAGATTGCCAGCTTTAAACTGACAGAAAGCGCCCTGCTGTTTT 337
Db 244 ACAAGATCTAGAGCAAAATAGCCCGCTTCCACGAGGGAGGATGCCATCTCTATC 303
Qy 338 CGAGCGGTACTTGGCCAAATGCGGTGCTTTTCATCCTTCCAGAAAGGAAGATGTCA 397
Db 304 CCAGCTGTTATGACGCCAAACCGCGCTCTTTGAGGCCCTGCTGACCCAGGAGCGCAG 363
Qy 398 TTTTAAGTGACAGCTCAATCATGCAAGTATGATGCGGCTGCGGACTTTCTAAGGCTG 457
Db 364 TCCTGTGCGAGCAGCTGAACCATGCTCCATCATGACGGCATCGGCTGTGCAAGGCC 423
Qy 458 ATACAGTTGTTTATCGGCATATTGATGAATGATCTTGAACCAAGCTGAATGAACAC 517
Db 424 ACAAGTACCGGTATGCGCACTTGACATGGCGACCTTAGAAGCCAAAGCTGAGAGGCC 483
Qy 518 AGCGTTATCAGCGCGCTTTTATCGGTAAACAGACGAGGATATTCAGCATGATGCGCAATCG 577
Db 484 AGAAGCATCGGCTGCGCCTGCTGCGCACTGATGGGCCCTTTTCCATGATGCGGACATCG 543
Qy 578 CCCCTCTTGATCAGATCATCTCACTTGGAAACGCTATCATGCTTGGTGTGCTTGATG 637
Db 544 CACCCCTGCGAGGAGATCTGCTGCGCTCTAGATATGCTGCGCTGCTCTTCATGGATG 603
Qy 638 ATGCCACGCAACAGGAGTTTTTGGCGGATTTCGGACAAAGGACGAGTGAATCTTTGGTG 697
Db 604 AATGCCATGCCACTGGCTTCTTGGGGCCACAGAGCGGGGACAGATGAGTGTGGTG 663
Qy 698 TTTGTCCCGACATGTTTATCGGCACCTTAAAGCAAGCTGTTGGCGCGGAAGAGGTTTTG 757
Db 664 TGATGGACCGGTCAACCATCATCAACTCCACCCTGGGGAGGCCCGGGTGGAGCATCAG 723
Qy 758 CGG 760
Db 724 GGG 726
```

## RESULT 10

```
CA778022
LOCUS ip16h09.x1 HR85 islet Homo sapiens cDNA clone IMAGE:6217600 3'
DEFINITION similar to SW:KBL HUMAN O75600 2-AMINO-3-KETOBUVRATE COENZYME A
LIGASE, MITOCHONDRIAL PRECURSOR ;, mRNA sequence.
```

CA778022

CA778022.1 GI:26015897

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominidae; Homo.

1 (bases 1 to 810)
Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemisha,I., Scarce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J.,
Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarisvilli,R.,
Williams,T., Jackson,Y. and Bowers,Y.

Endocrine Pancreas Consortium

Unpublished (2000)

Contact: Douglas Melton, Klaus H. Kaestner, &amp; Hiroshi Inoue

Endocrine Pancreas Consortium

Harvard University, Howard Hughes Medical Institute

Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,

MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Dr. Hiroshi Inoue
(hinoue@im.wustl.edu)

Possible reversed clone: similarity on wrong strand  
Seq primer: -40UP from Gibco  
High quality sequence stop: 439.

## FEATURES

## source

1. 810  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6217600"  
/tissue\_type="Purified pancreatic islet"  
/lab\_host="DH10B"  
/clone\_lib="HR85 islet"  
/notes="Organ: Pancreas; Vector: pBluescript SK(-); Site 1:  
NotI; Site 2: XhoI; cDNA made by oligo-dT priming.  
Size-selected on agarose gel. Average insert size ~1kb. 5'  
XhoI site was destroyed after directional cloning.  
Amplified once. Contact information: Hiroshi Inoue, MD,  
Metabolism Div. (Alan Permutt Lab), Washington University  
School of Medicine, Box 8127, 660 South Euclid Ave., St.  
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:  
314-362-1916, Fax: 314-747-2692."

## ORIGIN

Query Match 5.1%; Score 159.8; DB 6; Length 810;  
Best Local Similarity 53.2%; Pred. No. 8.9e-31;  
Matches 366; Conservative 0; Mismatches 313; Indels 9; Gaps 1;

Qy 158 CAACAATATTTAGGCTCGCAAGCATAGAGCTTTGATGATGCGCCCAACAGCAT 217  
Db 122 CCAACAACTACTCGGCTGAGCAGCACCCCTGAGGTGATCCAGCGAGGTCTGAGGCTC 181  
Qy 218 TGCAGCAATTTGGACAGGAAGCAGCGTTTACGTTTAAACGACAGGCAATTCGGTCTGGC 277  
Db 182 TGGAGGAGTTTGGAGCTGGCTCAGCTCTGCTCGCTTTTCTGTGGAAACCCAGAGCATCC 241  
Qy 278 ATGAAAAGCTAGAAAAGAAGATTGCCAGCTTTAACTGACAGAAAGCGCCCTGCTGTTTT 337  
Db 242 ACAAGATCTAGAAGCAAAATAGCCCGCTTCCACAGCGGAGGATGCCATCTCTATC 301  
Qy 338 CGAGCGGTACTTGGCCAAATGTCGGTGTCTTTTCATCTTGCACAGAAAGGAAGATGTCA 397  
Db 302 CCAGCTGTTATGACGCCAACCGCGCTCTTTGAGGCCCTGCTGACCCAGAGAGCGCAG 361  
Qy 398 TTTTAAGTGACAGCTCAATCATGCAAGTATGATGCGAGCGGTGCGGACTTCTTAAAGGTG 457  
Db 362 TCCTGTCGAGCAGCTGAAACCATGCTCCATCATCGACGGCATCGGCTGTGCAAGGCC 421  
Qy 458 ATACAGTTGTTTATCGGCATATTGATATGATGATTTGAAACCAAGCTGAATGAACAC 517  
Db 422 ACAAGTACCGCTATCGCCACCTGGACATGGCCGACCTAGAAGCCAAAGCTGCGAGGCC 481  
Qy 518 AGCGTTATCAGCGCGCTTTTATCGTAACAGACGAGGATTTTCAGCATGATGGCACAATCG 577  
Db 482 AGAAGCATCGGCTGCGCTGCTGGCCACTGATGGGGCTTTTTCATGATGGCAGCATCG 541  
Qy 578 CCCCTCTTGATCAGATCATCTCACTTCGAAACCTATCATGCTTCTGCTGCTGTTGATG 637  
Db 542 CACCCCTGACAGGAGATCTGCTGCTCGCTCTAGATATGGTCCCTGCTGCTTCTCATGATG 601  
Qy 638 ATGCCCAACGCAAGGAGTTTTGGCGGATTCGGGACAAAGGAACGAGTGAATCTTTGGTG 697  
Db 602 AATGCCATGCCACTGGCTTCTTGGGGCCCAACAGGACGGGGCACAGATGAGCTGTGGGTG 661  
Qy 698 TTTGTCCCGACATGTTATCGGCACCTTAAGCAAG-----CTGTTGGCGCGGAG 748  
Db 662 TGATGGACCAAGGTACCATCATCAACTCCACCTTGGGGAAGGCCCTGGGTGGAGCATCAG 721  
Qy 749 GAGGTTTTGCGGAGGATCAGCGTCTTCATCGACTTTTTCGTGAACCATGCCAGAACAT 808  
Db 722 GGGGCTACAGCAGGCGCTGGGCGCTTGGTGTGCTGCTGCGGACGGCGCCCGGCAT 781  
Qy 809 TTATCTTTTAAACCGCTATTTCGCCAGC 836



[illegible]



```
Db 19 CCAACAACACTCTGGGCTGAGCAGCCACCCTGAGGTGATCCAGGCGAGGTCTGCAGGCTC 78
Qy 218 TGCAGCAATTTGGGACAGGAGCGGTTTACGCTTTAAACGACAGCAATTCGTCTGGC 277
Db 79 TGGAGCAGTTTGGAGCTGGCCCTAGCTCTGCTCCCTTTATCTGTGGAACCCAGAGCATCC 138
Qy 278 ATGAAAAGCTAGAAAAGAGATTGCCAGCTTTAAACTGACAGAGACCGCCCTGCTGTTTT 337
Db 139 ACAAGATCTAGAAAGCAAAATAGCCCGCTTCCACAGCGGAGGATGCCATCTCTATC 198
Qy 338 CGAGCGTTACTTGGCCAAATGTGGGTCTCTTTTCATCTCTTGCAGAAAAGAGAGATGCA 397
Db 199 CCAGCTGTTTGGCGCAACCGCGGCTCTTTTGGAGCCCTGCTGACCCCGAGAGACGAG 258
Qy 398 TTTTAAGTGACAGCTCAATCATGCAAGATGATGACGAGCGTGGCGACTTTCTAAGGCTG 457
Db 259 TCCTGTGGAGAGCTGAAACATGCCCTCCATCATGAGCGGATCCGGCTGTGCAAGGCC 318
Qy 458 ATACAGTTGTTTATCGGCATATTGATATGATGATCTTGAACCAAGCTGAATGAACAC 517
Db 319 ACAAGTACCGCTATCGCCACCTGGACATGGCGGACCTNNNGCCAAAGCTGCNNNNNNN 378
Qy 518 AGGTTATCAGCGCGCTTTTATCGTAACAGAGAGATTTTTCAGCATGGATGGCAATCG 577
Db 379 NNAAGCATCGGCTCGGCTGGTGGCCACTGATGGGCGCTTTTCCATGGATGGCGACATCG 438
Qy 578 CCCCTCTTGATCAGATCATCTCAGTTCGGAACCTATCATGCTCGTGGTGGTTCATG 637
Db 439 CACCCCTGAGGAGATCTGCGGCTCGCCTTAGATATGGCGGCTCGTCTGTGGATG 498
Qy 638 ATGCCACGCAACAGGAGTTTGGGCGATTTCGGGACAGGAAACGAGTGAATATCTTGGTG 697
Db 499 AATGCCATGCCACTGGCTTCTGGGGCCACAGAGACGNNNNNNNNNAGCTGCTGGTG 558
Qy 698 TTTGTCCCGACATGTTTATCGGCACTTAAGCAAG-----CTGTTGGCGCGAAG 748
Db 559 TGATGGACCAAGGTCAACATCATCAACTCCACCTGGGGGAAGGCGCTGGTGGAGCATCAG 618
Qy 749 GAGGTTTGGCGGAGGATCAGCGGTCTTCATCGACTTTTGTGTAACCATGCCAGACAT 808
Db 619 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 678
Qy 809 TTATCTTTTCAAAACCGCTATTTCGGCCAGCAGCTGTGGGCTGTCTCAGAGGCTTTCAACA 868
Db 679 ACCTCTTCTCAACAGTCTGCCACTGTGTGCTGTGGCGGCTTCCAGGCCCTAGATC 738
Qy 869 TCATTGAAGCCAGCAGGGAAGAAACGACAGCTTTTATTTTCTTATATCAGCATGATCAGAA 928
Db 739 TGCTGATGGGAGTAAACACCATNGTCCAGTCTATGGCTGCAAGACCCAGAGGTTCCGTA 798
Qy 929 CCAGTCTGAAGATATGGGTTATGTGGTGAAGGAGATCACACCCGATTTATTCCTGTAG 988
Db 799 GTAAGATGGAAGTCTGGCTTCACTATCTCGGAGGCGAGTCAACCCATCTGCCCCGTGA 858
Qy 989 TCATTGGCGATGCCCAATAAACCGTCTCTATTGCTGAAAAACTCAGCGGCAAGGGAATTT 1048
Db 859 TGCTGGGTGATGCCCGGCTGCTCGCATGGCGGATGACATGCTGAAGAGGACATCT 918
Qy 1049 ATGCTCTGCCATTCGGCGCGCAACCGTTGGCGCGGTGAAAGCCGATTCG 1100
Db 919 TTGTCTCGGCTTCAGCTACCCCGTGTCCCAAGGCGAAGGCCCGGATCCG 970
```

## RESULT 14

BG752866

LOCUS

DEFINITION BG752866 787 bp mRNA linear EST 15-MAY-2001  
602732509F1 NIH\_MGC\_43 Homo sapiens cDNA clone IMAGE:4875915 5',  
mRNA sequence.

ACCESSION BG752866

VERSION BG752866.1 GI:14063519

KEYWORDS EST.

SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

## REFERENCE

1 (bases 1 to 787)

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)

Tissue Procurement: ATCC

cDNA Library Preparation: Ling Hong/Rubin Laboratory

DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLML)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LCM1758 row: a column: 04

High quality sequence stop: 785.

## FEATURES

source

1..787

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:4875915"

/tissue\_type="normal pigmented retinal epithelium"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH\_MGC\_43"

/note="Organ: eye; Vector: pOTB7; Site 1: XhoI; Site 2:

EcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCACGAG(G). Library constructed by Ling Hong

in the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies).

Note: this is a NIH\_MGC Library. |"

## ORIGIN

Query Match 5.0%; Score 157.2; DB 2; Length 787;  
Best Local Similarity 54.0%; Pred. No. 4.3e-10;  
Matches 321; Conservative 0; Mismatches 273; Indels 0; Gaps 0;

```
Qy 158 CAAACAATATTATTAGGCTCGCAAGCGATAGAGTTTGATCGATGCAAGCCCAACAGCAT 217
Db 69 CCAACAATACCTTGGGCTCGAGCAGCCACCCTGAGGTGATCCAGGCGAGTCTGCAGGCTC 128
Qy 218 TGCAGCAATTTGGGACAGGAGCGGTTTCAAGTTTAAACGACAGCAATTCGTCTGGC 277
Db 129 TGGAGGAGTTTGGAGCTGGGCTCAGCTCTGTCGCTTTATCTGTGGAACCCAGAGCATCC 188
Qy 278 ATCAAAAGCTAGAAAAGAGATTGCCAGCTTTTAAACTGACAGAGCGGCGCTGCTGTTTT 337
Db 189 ACAAGATCTAGAAGCAAAATAGCCGCTTCCACGCGGAGGATGCCATCTCTATC 248
Qy 338 CGAGCGTTACTTGGGCAATGTGCGGTCTCTTTTCATCTCTTGCAGAAAAGAGATGCA 397
Db 249 CCAGCTGTTATGACGCCCAACCGCGCTCTTTGAGGCGCTCTGACCCAGAGACGAG 308
Qy 398 TTTTAAGTGACAGCTCAATCATGCAAGTATGATCGAGCGCTGCCGATTTCTTAAGCTG 457
Db 309 TCCTGTGGACGAGCTGAACCATGCTCCATCATCGACGCGATCCGGCTGTGCAAGGCC 368
Qy 458 ATACAGTTGTTTATCGGCATATTGATATGAATGATCTTGAACCAAGCTGAATGAACAC 517
Db 369 ACAAGTACCGCTATCGCCACTCGACATGGCGGACCTAGAGCCAGCTGCAGAGGCGCC 428
Qy 518 AGCGTTATCAGCGCGGTTTTTATCGTAACAGACGAGGATTTTTCAGCATGGATGGCAATCG 577
Db 429 AGAAGCATCGGCTCGGCTGGTGGCCACTGATGGGCGCTTTTCCATGGATGGCGACATCG 488
Qy 578 CCCTCTTGTATCAGATCATCTCACTTTCGGAACCGCTATCATGCTTGTGGTGTGATG 637
Db 489 CACCCCTGCAAGGAGATCTGCTGCTCGCCTCTAGATATGGTGCCCTGCTCTTCATGGATG 548
```

Qy	638	ATGCCACGCAACAGGAGTTTGGCGGATTCGGGACAGGACGAGTCAATCTTGGTG	697
	Db		
	549	AATGCCATGCGACTGGCTTCTCTGGGGCCACAGACGGGGCACAGATGAGTGTGGGTG	608
	Qy		
Db	698	TTTGTCCCGACATTTGTTATCGGCACTTTAAGCAAAAGCTGTGGGCGGGAAGGAG	751
	609	TGATGGACCAAGTCACTCACTCCACCTGGGGAGGCGCTGGGTGGAG	662
RESULT 15			
LOCUS			
DEFINITION			
JGI CAAR2120.fwd NIH XGC tropliv1 Xenopus tropicalis cDNA clone			
IMAGE:7734748 5', mRNA sequence.			
DNO19683			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
Xenopus tropicalis (western clawed frog)			
Xenopus tropicalis			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;			
Xenopodinae; Xenopus; Silurana.			
1 (bases 1 to 871)			
Richardson,P., Lucas,S., Rokhsar,D., Dettler,J.C., Ng,D.C.,			
Brokstein,P. and Lindquist,E.A.			
DOE Joint Genome Institute Xenopus tropicalis EST project			
Unpublished (2004)			
Other ESTs: JGI CAAR2120.rev			
Contact: Lindquist,E.A., Richardson,P.			
DOE Joint Genome Institute			
2800 Mitchell Drive, Walnut Creek, CA 94598, USA			
Tel: 925 296 5600			
Fax: 925 296 5710			
Email: cdna@jgi-psf.org			
Tissue Procurement: Robert M. Grainger			
cDNA Library Preparation: Bruce Blumberg Laboratory, University of			
California, Irvine			
DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov			
Clone Distribution: I.M.A.G.E. Consortium/LLNL:			
http://image.llnl.gov			
Naming Conventions: EST name is generated by the concatenation of			
the JGI Clone Id and the direction of sequencing. The suffix '.fwd'			
indicates a forward sequencing read of the insert. It does not			
necessarily reflect the orientation of the insert.			
Plate: CAAR 0021 row: o column: 2			
High quality sequence stop: 863.			
FEATURES	Location/Qualifiers		
	1..871		
source	/organism="Xenopus tropicalis"		
	/mol_type="mRNA"		
	/strain="N6 (Nigerian 6th generation inbred)"		
	/db_xref="taxon:8364"		
	/clone="IMAGE:7734748"		
	/sex="male"		
	/tissue_type="Liver"		
	/dev_stages="Adult"		
	/lab_host="ElectroMAX DH10B T1 Phage Resistant cells"		
	/clone_lib="NIH XGC tropliv1"		
	/notes="Vector: pCS107; Site 1: EcoRI; Site 2: XhoI; The		
	library was prepared from 5 ug of poly A+ RNA by oligo-dT		
	priming		
	(5'-GATGAGAGAGAGAGAGAGACTAGTCTCGAGTTTTTTTTTTTTTTT-3')		

Search completed: February 11, 2006, 23:19:32

Job time : 8022 secs

ORIGIN			
Query Match			
Best Local Similarity 53.9%; Score 157.2; DB 8; Length 871;			
Matches 371; Conservative 0; Mismatches 308; Indels 9; Gaps 2;			
Qy	158	CAACAATTTATTTAGGCTCGCAAGCGATAGAGCTTTTGATGATGAGCCCAACAGCAT	217
	Db		
Db	177	CCATATACTACTTAGGCTTTCCAGCCATCCCAAGGTGATCCGTCGTCGATGTGAGAC	236
	Qy		
Qy	218	TGCAGCAATTTGGGACAGGAGCGGTTTCACTTTAAACGACAGGCAATTCGCTTGGC	277
	Db		
Db	237	TAGAGGAGTTTGGAGCGGCTCTCAGCTCAGTGAGGTTTATCTGTGGGACACAGAATATCC	296
	Qy		
Qy	278	ATGAAAGCTAGAAAAGAAATTCGACAGCTTTAACTGACAGAACGCGCCCTGCTGTTT	337
	Db		
Db	297	ATAAGAGTCTAGAACAGAGATTTGCCCGGTTCCACAGAGAGAGATGCCATCTTGTATA	356
	Qy		
Qy	338	CGAGCGTTTACTTTGGCCCAATGTGCGTGTCTTTTCATCTTCCAGAAAAAGAGATGTCA	397
	Db		
Db	357	TCAGCTCTTTTGATGCCCAATGCTGGTATATTTGAGGCTCTTCTGTCCTGAAATGCAG	416
	Qy		
Qy	398	TTTTAAGTGACCAAGCTCAATCATGCAAGTATGATCGACGGCTGCCGACTTTTCTAAGGCTG	457
	Db		
Db	417	TGCTCTCAGATGAGTTGAACACCGCCTCAATCATTTGATGGTATCCGCTTTTGCAGGCCA	476
	Qy		
Qy	458	ATACAGTTGTTTATCGGCATATTGATATGAATGATCTTTGAAACAAGCTGAATGAACAC	517
	Db		
Db	477	ATAAGTACCGCTATAAGCACATGAGACCTAGATGATTTAGAAAGCCCAAGCTTTAAGAGGCAC	536
	Qy		
Qy	518	AGCGTTATACGCGCGCTTTTATGTTACACAGACGAGTATTTCAGCATGGATGGCACAAATCG	577
	Db		
Db	537	AGAAGCACCGCTTGGCTCTCATTTGCCACACAGCGGTGTTTTTCTATGACGAGACATTG	596
	Qy		
Qy	578	CCCTCTTGATCAGATCATCTCACTTGCAGAAAGCTATCATGCTTCTGTTGCTCGTTGATG	637
	Db		
Db	597	CACCTCTTCTGTGAGATTGCAACTTAGCAAAAGCGCTACAATGCCCTTAGTCTTTATAGATG	656
	Qy		
Qy	638	ATGCCACGCAACAGGAGTTTGGGCGATTCGGGACAGGAAAGACGAGTGAATACTTTGGTG	697
	Db		
Db	657	AGTGCCATGTCTACCGGTTTTTATGGGACCCAAATGGAAGGGGACAGATGAGCTTCTCGAG	716
	Qy		
Qy	698	T-----TTGTCCCGACATTGTTATCGGCACCTTTAAGCAAAAGC---TGTTGGCGCGGAAG	748
	Db		
Db	717	TCATGATCAAGTGACAAATTTGTAATTTACTACCTTGGCAAAAGCACTGGAGGGGCTGCAG	776
	Qy		
Qy	749	GAGGTTTGGCGGAGGATCAGCGGTCTTCATTCGACTTTTTTGTGAAACCATGCCAGAACAT	808
	Db		
Db	777	GTGGGTACACCACTGGGCCCAAGCCTTTGATTTGACTTACTACGGCAGCGCTCTCGGCCAT	836
	Qy		
Db	809	TTATCTTTCAACCGCTATTTCGCGCAGC	836
	837	ACCTGTTTCTCAAAACACCCCTCCGCGCAGC	864

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 11, 2006, 21:06:35 ; Search time 384 Seconds  
(without alignments)

14609.335 Million cell updates/sec

Title: US-10-681-086-1

Perfect score: 3156

Sequence: 1 ggatccacgaggttacgagc.....atcggtatgatgtcgaaattc 3156

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA.\*

- 1: /cgn2\_6/ptodata/1/ina/1 COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5 COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/H COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/PPUS COMB.seq.\*
- 7: /cgn2\_6/ptodata/1/ina/PP COMB.seq.\*
- 8: /cgn2\_6/ptodata/1/ina/RE COMB.seq.\*
- 9: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3156	100.0	3156	3	US-09-633-927-1
2	1612.6	51.1	8478	3	US-08-676-818-1
3	1612.6	51.1	8478	3	US-09-407-549-1
4	1612.6	51.1	8478	3	US-09-407-728-1
5	867	27.5	5793	3	US-09-869-855A-3
6	867	27.5	5943	3	US-09-869-855A-2
7	867	27.5	7336	3	US-08-418-085-3
8	867	27.5	7336	3	US-09-099-011A-3
9	867	27.5	7336	3	US-08-470-369-21
10	867	27.5	7336	3	US-08-098-877B-3
11	867	27.5	8119	2	US-08-460-343B-1
12	867	27.5	8119	2	US-08-398-028B-1
13	867	27.5	8119	2	US-08-504-265B-1
14	867	27.5	9144	3	US-08-556-978B-79
15	865.4	27.4	7326	9	5304637-21
16	853	27.0	5609	3	US-09-313-677-14
17	853	27.0	5609	3	US-09-313-677-15
18	853	27.0	7026	3	US-09-313-677-20
19	853	27.0	7344	3	US-09-313-677-16
20	762.4	24.2	1200	2	US-08-642-045B-3
21	762.4	24.2	1200	2	US-08-852-268-1
22	759	24.0	759	3	US-09-697-186B-10
23	703	22.3	6169	2	US-08-875-154-2
24	701	22.2	6837	3	US-09-928-847B-49

#### ALIGNMENTS

##### RESULT 1

US-09-633-927-1  
; Sequence 1, Application US/09633927  
; Patent No. 6656721  
; GENERAL INFORMATION:  
; APPLICANT: HOHMANN DR., Hans-Peter  
; APPLICANT: MOUNCEY DR., Nigel J.  
; APPLICANT: SCHLIEKER DR., Heinrich W.  
; APPLICANT: STEBBINS DR., Jeffrey W.  
; TITLE OF INVENTION: PROCESS FOR PRODUCING A TARGET FERMENTATION PRODUCT  
; FILE REFERENCE: Process For Prod. Target Ferm. Product  
; CURRENT APPLICATION NUMBER: US/09/633,927  
; CURRENT FILING DATE: 2000-08-08  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 3156  
; TYPE: DNA  
; ORGANISM: Bacillus subtilis  
US-09-633-927-1

Query Match	100.0%;	Score 3156;	DB 3;	Length 3156;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3156;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GGATCCACGAGGTTACGAGCCTTGAAGATTGATTCCTGTTAAACGAGCGGTTAGACAGA	60	
Db	1	GGATCCACGAGGTTACGAGCCTTGAAGATTGATTCCTGTTAAACGAGCGGTTAGACAGA	60	
Qy	61	ATGAAGAAGCCGCGGTACATCGTAACCTCGGTCAATGGATGGAGCCGCGGTTCCAGAG	120	
Db	61	ATGAAGAAGCCGCGGTACATCGTAACCTCGGTCAATGGATGGAGCCGCGGTTCCAGAG	120	
Qy	121	AGGAATTTGATGGCGAAATCAACGGTCTGTCCTCAACCAATTTATTAGGCTCGCA	180	
Db	121	AGGAATTTGATGGCGAAATCAACGGTCTGTCCTCAACCAATTTATTAGGCTCGCA	180	
Qy	181	AGCGATAGAGTTTGTATCGATCGAGCCCAACAGCAATTCGAGCAATTTGGGACAGGAGC	240	
Db	181	AGCGATAGAGTTTGTATCGATCGAGCCCAACAGCAATTCGAGCAATTTGGGACAGGAGC	240	
Qy	241	AGCGGTTTACGAGCAATTCGGTCTGGCATGAAAGCTAGAAAAGAGATT	300	
Db	241	AGCGGTTTACGAGCAATTCGGTCTGGCATGAAAGCTAGAAAAGAGATT	300	
Qy	301	GCACGCTTTAACTGACAGAGCGGCCCTGCTGTTTCGAGCGGTTACTTGGCCCAATGTC	360	
Db	301	GCACGCTTTAACTGACAGAGCGGCCCTGCTGTTTCGAGCGGTTACTTGGCCCAATGTC	360	

Sequence 1, Appli  
Sequence 5, Appli  
Sequence 3541, Ap  
Sequence 3934, Ap  
Sequence 3895, Ap  
Sequence 634, App  
Sequence 539, App  
Sequence 2561, Ap  
Sequence 7, Appli  
Sequence 7, Appli  
Sequence 7, Appli  
Sequence 5, Appli  
Sequence 5, Appli  
Sequence 5, Appli  
Sequence 338, App  
Sequence 338, App  
Sequence 110, App  
Sequence 1, Appli  
Sequence 1, Appli

Db 301 GCCAGCTTTAAACTGACAGAAAGCGCCCTGCTGTTGTTTTTCGAGCGGTTACTTTGGCCAAATGTC 360  
Qy 361 GGTGTCCTTTTCATCTCTGCGAGAAAAGGAAGATGTCATTTTAAAGTGAACGAGCTCATATCAT 420  
Db 361 GGTGTCCTTTTCATCTCTGCGAGAAAAGGAAGATGTCATTTTAAAGTGAACGAGCTCAATCAT 420  
Qy 421 GCAAGTATGATCGAGCGGTGCGGACTTTCTAAGGCTGATACAGTTGTTTATCGCATATT 480  
Db 421 GCAAGTATGATCGAGCGGTGCGGACTTTCTAAGGCTGATACAGTTGTTTATCGGCAATT 480  
Qy 481 GATATGAATGATCTTTGAAAACAAGCTGAATGAAAACACAGCGTTATCAGCGCGCTTTTATC 540  
Db 481 GATATGAATGATCTTTGAAAACAAGCTGAATGAAAACACAGCGTTATCAGCGCGCTTTTATC 540  
Qy 541 GTAACAGACGGAGATTTACAGCATGATGGGCACATCGCCCTCTTTGATCAGATCATCTCA 600  
Db 541 GTAACAGACGGAGATTTACAGCATGATGGGCACATCGCCCTCTTTGATCAGATCATCTCA 600  
Qy 601 CTTGCGAAACGCTATCATGCCCTTCGTGCTCGTTGATGATGCCACGCAACAGGAGTTTTG 660  
Db 601 CTTGCGAAACGCTATCATGCCCTTCGTGCTCGTTGATGATGCCACGCAACAGGAGTTTTG 660  
Qy 661 GCGGATTCGGGACAAAGAACGAGTGAATPACTTTGGTGTGTTGTCGCCGACATTTGTTATCGGC 720  
Db 661 GCGGATTCGGGACAAAGAACGAGTGAATPACTTTGGTGTGTTGTCGCCGACATTTGTTATCGGC 720  
Qy 721 ACCTTAAGCAAAAGCTGTTGGCGCGGAGGAGTTTTGGCGGAGGATCAGCGGTCTTCATC 780  
Db 721 ACCTTAAGCAAAAGCTGTTGGCGCGGAGGAGTTTTGGCGGAGGATCAGCGGTCTTCATC 780  
Qy 781 GACTTTTTGCTGAACCATGCCAGAACATTTATCTTTCAAACCGCTATTTCGCCAGCCAGC 840  
Db 781 GACTTTTTGCTGAACCATGCCAGAACATTTATCTTTCAAACCGCTATTTCGCCAGCCAGC 840  
Qy 841 TGTGCGGCTGCTACGAGGCTTTCAACATCATTTGAAGCCAGCAGCGGAAAAACGACAGCTT 900  
Db 841 TGTGCGGCTGCTACGAGGCTTTCAACATCATTTGAAGCCAGCAGCGGAAAAACGACAGCTT 900  
Qy 901 TTATTTTCTTATATCAGCATGATCAGAACCCAGTCTGAAGAAATATGGGTATATGGTGTGAAA 960  
Db 901 TTATTTTCTTATATCAGCATGATCAGAACCCAGTCTGAAGAAATATGGGTATATGGTGTGAAA 960  
Qy 961 GGAGATCACACACCGATTAATCTGTAGTCAATGGCGATGCCATATAAACCGTCTCTATTT 1020  
Db 961 GGAGATCACACACCGATTAATCTGTAGTCAATGGCGATGCCATATAAACCGTCTCTATTT 1020  
Qy 1021 GCTGAAAACTGCGAGGCAAGGGAATTTATGCTCTCTGCAATTCGGCCGCCAACCGTTGCG 1080  
Db 1021 GCTGAAAACTGCGAGGCAAGGGAATTTATGCTCTCTGCAATTCGGCCGCCAACCGTTGCG 1080  
Qy 1081 CCGGTTGAACCGCGATTGGAAGCTTGGGACGACAGTTCGAGATCAGGGAATGAGTTTATA 1140  
Db 1081 CCGGTTGAACCGCGATTGGAAGCTTGGGACGACAGTTCGAGATCAGGGAATGAGTTTATA 1140  
Qy 1141 AAATAAAAAAGCACCTGAAAAGGTGCTTTTTTTTGTAGTTGTTGAACTGCTCTCTTTCTT 1200  
Db 1141 AAATAAAAAAGCACCTGAAAAGGTGCTTTTTTTTGTAGTTGTTGAACTGCTCTCTTTCTT 1200  
Qy 1201 ATCTTGATACATATAGAAATAACGTCATTTTATTTTATTTTATTTAGTTGCTGAAAAGGTGCG 1260  
Db 1201 ATCTTGATACATATAGAAATAACGTCATTTTATTTTATTTTATTTAGTTGCTGAAAAGGTGCG 1260  
Qy 1261 TTGAAGTGTGGTATGATGTTTTTAAAGTATTTGAAACCCCTTAAATTTGTTGTCACAG 1320  
Db 1261 TTGAAGTGTGGTATGATGTTTTTAAAGTATTTGAAACCCCTTAAATTTGTTGTCACAG 1320  
Qy 1321 AAAAAACCCATCTGTATAAGTTATAAGTGACTAAACCAATAACTAAATAGATGGGGTTT 1380  
Db 1321 AAAAAACCCATCTGTATAAGTTATAAGTGACTAAACCAATAACTAAATAGATGGGGTTT 1380  
Qy 1381 CTTTAAATATTATGTGCTTAATAGTAGCACTTTATTCAGATGAAAAATCAAGGGTTTTAG 1440  
Db 1381 CTTTAAATATTATGTGCTTAATAGTAGCACTTTATTCAGATGAAAAATCAAGGGTTTTAG 1440

Qy 1441 TGGACAGACAAAAGTGGAAAAGTGAACAATGCTGTAGGAAGACGAGTTATTAATAG 1500  
Db 1441 TGGACAGACAAAAGTGGAAAAGTGAACAATGCTGTAGGAAGACGAGTTATTAATAG 1500  
Qy 1501 CTGAATAAGAACGGTGTCTCCAAATATTCTTATTTAGAAAAGCAAACTCTAAAAATTATCT 1560  
Db 1501 CTGAATAAGAACGGTGTCTCCAAATATTCTTATTTAGAAAAGCAAACTCTAAAAATTATCT 1560  
Qy 1561 GAAAAAGGAATGAGAATAGTGAATGGACCAATAATAATAGCTAGAGAAGAAGAAATGAAG 1620  
Db 1561 GAAAAAGGAATGAGAATAGTGAATGGACCAATAATAATAGCTAGAGAAGAAGAAATGAAG 1620  
Qy 1621 ATTGTTTCATGAATAAGGAACGAATATTGGATAAATATGGGGATGATGTTAAAGCTATT 1680  
Db 1621 ATTGTTTCATGAATAAGGAACGAATATTGGATAAATATGGGGATGATGTTAAAGCTATT 1680  
Qy 1681 GGTGTTTATGGCTCTCTTGGTCTGTCAGACTGATGGGCCCTATTTCGGGATTTGAGATGATG 1740  
Db 1681 GGTGTTTATGGCTCTCTTGGTCTGTCAGACTGATGGGCCCTATTTCGGGATTTGAGATGATG 1740  
Qy 1741 TGTGTCATGTCACACAGAGGAAGCAGAGTTTCAGCCATGAATGGACAACCGGTGAGTGGAAAG 1800  
Db 1741 TGTGTCATGTCACACAGAGGAAGCAGAGTTTCAGCCATGAATGGACAACCGGTGAGTGGAAAG 1800  
Qy 1801 GTGGAAGTGAATTTTGATAGCGAAGAGATTTCTACTAGATTATGATCTCTCAGGTGGAATCA 1860  
Db 1801 GTGGAAGTGAATTTTGATAGCGAAGAGATTTCTACTAGATTATGATCTCTCAGGTGGAATCA 1860  
Qy 1861 GATTGGCCGCTTACACATGTTCAATTTTTCTTATTTTGGCGATTTTATGATTCAGGTGGA 1920  
Db 1861 GATTGGCCGCTTACACATGTTCAATTTTTCTTATTTTGGCGATTTTATGATTCAGGTGGA 1920  
Qy 1921 TACTTAGAAGAAAGTATCAAACTGCTAAATCGGTAGAGGCCAAACGTTCCACGATGCG 1980  
Db 1921 TACTTAGAAGAAAGTATCAAACTGCTAAATCGGTAGAGGCCAAACGTTCCACGATGCG 1980  
Qy 1981 ATTTGTCGCCCTTATCGTAGAAGAGCTGTTTGAATATCGACGCAATATGCGTAATATTCGT 2040  
Db 1981 ATTTGTCGCCCTTATCGTAGAAGAGCTGTTTGAATATCGACGCAATATGCGTAATATTCGT 2040  
Qy 2041 GTGCAAGACGCGCAACAACTTTTACCATCTCTGATGTCAGAGTAGCAATGGCAGGTGCC 2100  
Db 2041 GTGCAAGACGCGCAACAACTTTTACCATCTCTGATGTCAGAGTAGCAATGGCAGGTGCC 2100  
Qy 2101 ATGTTGATTTGGTCTGCAATCATCGCATCTGTTTATACGACGAGCGCTTCGGTCTTAACTGAA 2160  
Db 2101 ATGTTGATTTGGTCTGCAATCATCGCATCTGTTTATACGACGAGCGCTTCGGTCTTAACTGAA 2160  
Qy 2161 GCAGTTAAGCAATCAGATCTTCTTCAGGTTTATGACCATCTGTGCCAGTTCTGTAATGTCT 2220  
Db 2161 GCAGTTAAGCAATCAGATCTTCTTCAGGTTTATGACCATCTGTGCCAGTTCTGTAATGTCT 2220  
Qy 2221 GGTCAACTTTTCCGACTCTGAGAAAATTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATT 2280  
Db 2221 GGTCAACTTTTCCGACTCTGAGAAAATTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATT 2280  
Qy 2281 CAGAGTGGACAGAAACGACACGGATATATAGTGGATGTGTCAAAACGCAATACCAATTTTGA 2340  
Db 2281 CAGAGTGGACAGAAACGACACGGATATATAGTGGATGTGTCAAAACGCAATACCAATTTTGA 2340  
Qy 2341 ATTGGAAGCGCGGATTTGAGTCTTTACCGGATGGTGAATAAGGAAACGCTGCTTGAAGCGG 2400  
Db 2341 ATTGGAAGCGCGGATTTGAGTCTTTACCGGATGGTGAATAAGGAAACGCTGCTTGAAGCGG 2400  
Qy 2401 CGAAGCGGGCGACGATCTGAAATATCGGCACATATTTGATCGTGGCAAGCGGACGAGGTC 2460  
Db 2401 CGAAGCGGGCGACGATCTGAAATATCGGCACATATTTGATCGTGGCAAGCGGACGAGGTC 2460  
Qy 2461 CGTCTAACAGAGAAGTGGATCAGGTCTAGATCGGTTTCAGGAAATTTAAAGACGCTATG 2520  
Db 2461 CGTCTAACAGAGAAGTGGATCAGGTCTAGATCGGTTTCAGGAAATTTAAAGACGCTATG 2520

Qy	2521	GACTGAAGATTGTGCATGCTCTTGGACTGTTGAAGCCAGAGACGCGGAGCGGCTCAAAG	2580
Db	2521	GACTGAAGATTGTGCATGCTCTTGGACTGTTGAAGCCAGAGACGCGGAGCGGCTCAAAG	2580
Qy	2581	ATGCAGGAGTAGACCGCTATAATCATAAATTTGAATAGCTCACAGAGAAAACATTCAAACA	2640
Db	2581	ATGCAGGAGTAGACCGCTATAATCATAAATTTGAATAGCTCACAGAGAAAACATTCAAACA	2640
Qy	2641	TCACAACCTCACATACATACGATGACAGAGTCAATACGGTTGAAATCGCAAAAGAAATCGG	2700
Db	2641	TCACAACCTCACATACATACGATGACAGAGTCAATACGGTTGAAATCGCAAAAGAAATCGG	2700
Qy	2701	GGCTGCTCTCGTGTTTCAAGCGCCATTATCGGGATGGAAGGAGAGGAAACACGATGTCAATG	2760
Db	2701	GGCTGCTCTCGTGTTTCAAGCGCCATTATCGGGATGGAAGGAGAGGAAACACGATGTCAATG	2760
Qy	2761	ACATCGCCAAAAGCTTGAAGGCTCTTGACCGGGAATCCATTCTCTGTGAAATTTTTTGCATG	2820
Db	2761	ACATCGCCAAAAGCTTGAAGGCTCTTGACCGGGAATCCATTCTCTGTGAAATTTTTTGCATG	2820
Qy	2821	CAATTGATGSCACGCGGTTAGAGGCGTCAACGNAATTAACCCGCGCTGTATGTTTAAAG	2880
Db	2821	CAATTGATGSCACGCGGTTAGAGGCGTCAACGNAATTAACCCGCGCTGTATGTTTAAAG	2880
Qy	2881	TGCTGGCGCTGTTTCCGTTTTATCAATCCATCAAAAGAAATTCGCATTTCCGGAGGAAGAG	2940
Db	2881	TGCTGGCGCTGTTTCCGTTTTATCAATCCATCAAAAGAAATTCGCATTTCCGGAGGAAGAG	2940
Qy	2941	AGGTCAATCTCCGCACATTCAGCCATTAGGGCTTTACGCGCAAACTCCATTTTTGTGTCG	3000
Db	2941	AGGTCAATCTCCGCACATTCAGCCATTAGGGCTTTACGCGCGCAAACTCCATTTTTGTGTCG	3000
Qy	3001	GAGACTACTTAAACAACTTCGCGGCAAGAGAGACGGAGGATCATATAATGCTGAGTGATT	3060
Db	3001	GAGACTACTTAAACAACTTCGCGGCAAGAGAGACGGAGGATCATATAATGCTGAGTGATT	3060
Qy	3061	TAGGCTTTGAAGTTGAAATCAGTCCAGAAATTAAGGCTAGTTTTAAGTGCAGAAAGCTGAA	3120
Db	3061	TAGGCTTTGAAGTTGAAATCAGTCCAGAAATTAAGGCTAGTTTTAAGTGCAGAAAGCTGAA	3120
Qy	3121	AGAAATCAATAAAAGCAATCGGTATGATGTGCAATTC	3156
Db	3121	AGAAATCAATAAAAGCAATCGGTATGATGTGCAATTC	3156

## RESULT 2

```

US-08-676-818-1
; Sequence 1, Application US/08676818
; Patent No. 6057136
; GENERAL INFORMATION:
; APPLICANT: Bower, Stanley Grant
; APPLICANT: Perkins, John B.
; APPLICANT: Yocum, R. Rogers
; APPLICANT: Peto, Janice G.
; TITLE OF INVENTION: BIOTIN BIOSYNTHESIS IN BACILLUS
; TITLE OF INVENTION: SUBTILIS
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: Wordperfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/676,818
; FILING DATE: 08-JUL-1996
; CLASSIFICATION:

```



Qy 722 CCTTAAGCAAAAGCTGTTGGCGGGAAGGAGGTTTGGCGGAGGATCGAGCGTCTTCATCG 781  
Db CTTTAAGCAAAAGCTGTTGGCGGGAAGGAGGTTTGGCGGAGGATCGAGCGTCTTCATCG 3303  
Qy 782 ACTTTTGTCTGAACCATGCGCAGAACATTTATCTTTCAAACCGCTATTCCGCGCACCGAGCT 841  
Db ACTTTTGTCTGAACCATGCGCAGAACATTTATCTTTCAAACCGCTATTCCGCGCACCGAGCT 3363  
Qy 842 GTGCGGCTGCTCAGAGGCTTTCAACATCATTTGAAGCCAGCAGGNAAGAACGACAGCTTT 901  
Db GTGCGGCTGCTCAGAGGCTTTCAACATCATTTGAAGCCAGCAGGNAAGAACGACAGCTTT 3423  
Qy 902 TATTTTCTTATATCAGCATGATCAGAACCAAGTCTGAAGAAATATGGGTTATGTGTGAAG 961  
Db TATTTTCTTATATCAGCATGATCAGAACCAAGTCTGAAGAAATATGGGTTATGTGTGAAG 3483  
Qy 962 GAGATCACACACCGGATTTCTGTAGTCATTTGGCGATGCCATATAAACCGTCTCTATTTG 1021  
Db GAGATCACACACCGGATTTCTGTAGTCATTTGGCGATGCCATATAAACCGTCTCTATTTG 3543  
Qy 1022 CTGAAGAACTCGAGGCGAAGGAAATTTATGCTCTGCCATTTGGCGCGCCACCGTTGGC 1081  
Db CTGAAGAACTCGAGGCGAAGGAAATTTATGCTCTGCCATTTGGCGCGCCACCGTTGGC 3603  
Qy 1082 CGGGTGAAGCCGGATTCGAAGCTTGGGCGACAGTCTGAGATCAGGGAAATGAGTTTATAA 1141  
Db CGGGTGAAGCCGGATTCGA-----ATTA 3627  
Qy 1142 AATAAAAAAGCACTGAAAAGGTGCTTTTTTTTGTAGTGTGTTTGAACCTGTTCTTTA 1201  
Db CAATCAGCTCTGACCAACAGTATGGGTGATATGATCATTTGCTGCAAAACATTTCAITCAA 3687  
Qy 1202 TCTTGATACATATAGAAATACGTCATTTTATTTTATTTTATTTTGTCTGTAAGGTGCGT 1261  
Db TCGAAAGAGGAGCTGCACATCATTTGAGGGGTTTTTTTGTGACGGGAACCTGA----- 3738  
Qy 1262 TGAAGTGTGTATGTATGCTTTTAAAGTATGAAAACCCCTTAAATTTGGTTGCACAGA 1321  
Db -----TACAGAAGTAGGNAACCGTTATATCCAGCGTCTTG- 3776  
Qy 1322 AAAACCCCATCTGTTAAAGTTTATAAGTGACTAAACAAATAACTAAATAGATGGGGTTTC 1381  
Db -----CTGCTTATTGAAAGACAAATAATAGACATGTGCGGGGTGA 3816  
Qy 1382 TTTTAAATATATGTCCTCTATAGTATGACATTTATTCAGATGAAAATCAAGGTTTATG 1441  
Db TAAACCATTTTAAAGCGGATATCGGCCATCATCCAGATAGTATACAAAGTTTGTGAA 3876  
Qy 1442 GGACAAGACAAAAAGTGGAAAAGTGAGACCATGTGCTTTAGGAAGACGAGTTATTAAATAGC 1501  
Db AGATATGTCGAGA----- 3890  
Qy 1502 TGAATAAGACGGTCTCTCAAATATTTCTTTATTTAGAAAAGCAAAATCTAAAATTTATCTG 1561  
Db -----CCAGTCTTCTCATGAAGACATTACGCTTTTGGCTTCAA 3930  
Qy 1562 AAAAGGNAAGCAATAGTGAATGGAACCAATAATATGACTAGAGAAGAAAGATGAAGA 1621  
Db GGCCCGCTTGCAACATACATGTTGCGGGAACCTTTGAGGAAAAGACTGTCAACATGGAAGA 3990  
Qy 1622 TTGTTTCATGAAATTAAGGAACGAATATTTGGAATAAATATGGGATGATGTTAAGGCTATTG 1681  
Db GGTTTTAAAGCCATTTGGGGCGGATAGAGAAAA-----CATGAATGCTTCATCGTAG 4043  
Qy 1682 GTGTTTATGCTCTTGTGCTGAGACTGATGGGCCCTATTTCCGATATTGAGATGATGT 1741  
Db AAGGTGACGGCGGTATTCTGTGCCATTTGGGAGAGGACTATTTGGTCAGTCATGTCATAA 4103  
Qy 1742 GTGTCATGTCAACAGAGGAACGAGTTCAGCCATGAATGGACAACCGGTGAGTGGGAAG 1801  
Db -----AAGCGTTGACGCTTCCCATGATTTATGTGGCGCTCTCGCTGCTGGAAAC 4153  
Qy 1802 TGGAAAGTAAATTTTGATAGCAAGAGATTTCTACTAGATTTATGATCTCAGGTGGAATCAG 1861

Db 4154 ATTAATCATACCTTTTAACTGTCAAATATGCAAGAAAGCATGGGCTTCCAATCGCGGA 4213  
Qy 1862 ATTGGCCCGCTTACACATGGTCAATTTTCTCTATTTTGGCCGATTTATGATTCAGGTGGAT 1921  
Db 4214 ATTATCATCAATGGAATCAGTGACTCTCTCTGATGAAGATGAAAAACCAATCTCTGAGATG 4273  
Qy 1922 ACTTAGAGAAAGTGTATCAAACTGCTAAATCGGTAGAGGCCAAACGTTCCACATGCGA 1981  
Db 4274 ATTGAGGCTTATGCGGTGCGGATTTTAGGGTTAGCCCAAGCTTGCACACGTG--- 4330  
Qy 1982 TTTGTGCCCTTATCGTAGAAGAGCTGTTTGAATATCAGGCAAAATGCGCTAATATTCGTG 2041  
Db 4331 -----ACGAAAGAAACGGTTCTACATATGTTAAAGACCAT----- 4367  
Qy 2042 TGCAGGACCGACAAACATTTCTACCATCTTGACTGTACAGGTAGCAATGGCAGGTGCCA 2101  
Db 4368 -----TCNATCTATCTACTGATGAATCAAGTGGGGTATGA 4405  
Qy 2102 TGTGTATTGCTGTCATCGCATCTGTTATACGACGAGCGCTTCGGTCTTAACTGAAG 2161  
Db 4406 GAATGAATCAATGATGGAACCTGCAGACCGGGTCTGGCTGGAGCAAGAGTGACTGACG 4465  
Qy 2162 CAGTTAAGCAATCAGATCTTCTTTCAGGTTATGACCATCTGTGCCAGTTCGTAAATGCTG 2221  
Db 4466 AAGAGGCGCTTCAATATTACATTTGCTCTGATGAAGATATTTTCTGCTATTAATGACACGGG 4525  
Qy 2222 GTCMACTTTCCGACTCTGAGAAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTC 2281  
Db 4526 CTTTTCATCATCAGAAAAACCTTTTACGNAAAAAAGTAAAGCTCAATATGATTATGAATG 4585  
Qy 2282 AGAGTGGACAGAACGACACGGATATATAGTGGATGTCAAACGCGATACCAATTTTGAA 2341  
Db 4586 CGAATCCGGGCTCTGCCCGAAAACTGCGGCTATTGTTCA-----GTCTGCGAT 4637  
Qy 2342 TTCGAAACCGCCGATTGAGTCTTACCGGATGGTGAATTAAGAAACCGTCTGTAAGGCGC 2401  
Db 4638 TTCGAAACCGCCGATTGAGTCTTACCGGATGGTGAATTAAGAAACCGTCTGTAAGGCGC 4697  
Qy 2402 GAAGCGGCGCACGATCTGAATATCGGCACATATTGATCGTGGCAAGCGGACAGGTCC 2461  
Db 4698 GAAGCGGCGCACGATCTGAATATCGGCACATATTGATCGTGGCAAGCGGACAGGTCC 4757  
Qy 2462 GTCTAACAGAGAAAGTGGATCAGGTCTGATGCGGTTTCAAGAAATTAAGAGAGAGTATGG 2521  
Db 4758 GTCTAACAGAGAAAGTGGATCAGGTCTGATGATGCGGTTCAGGNAATTAAGAGAGAGTATGG 4817  
Qy 2522 ACTGAAGATTTGTGCAATGCTTTGGAAGTGTGGAAGCCAGAGAGCGGCGGCTCAAGA 2581  
Db 4818 ACTGAAGATTTGTGCAATGCTTTGGAAGTGTGGAAGCCAGAGAGCGGCGGCTCAAGA 4877  
Qy 2582 TGCAGGATGACCGCTTATTAATCATTAATTTGAATACGTCAAGAGAAACCAATTCACACAT 2641  
Db 4878 TGCAGGATGACCGCTTATTAATCATTAATTTGAATACGTCAAGAGAAACCAATTCACACAT 4937  
Qy 2642 CACAACTCATACATACATACGATGACAGTCAATACGTTTGAATCGCAAGAGAAAGTATCGG 2701  
Db 4938 CACAACTCATACATACATACGATGACAGTCAATACGTTTGAATCGCAAGAGAAAGTATCGG 4997  
Qy 2702 GCTGTCTCCGTGTTCCAGGCGCCATTATCGGATGAAGAGACGAAACAGGATGCTATTGA 2761  
Db 4998 GCTGTCTCCGTGTTCCAGGCGCCATTATCGGATGAAGAGACGAAACAGGATGCTATTGA 5057  
Qy 2762 CATCGCAAAAGCTTGAAGGCTCTTGACGCGGATTCCTCTCTGTGAATTTTTTTCATGTC 2821  
Db 5058 CATCGCAAAAGCTTGAAGGCTCTTGACGCGGATTCCTCTCTGTGAATTTTTTTCATGTC 5117  
Qy 2822 AATTGATGGCAGCGGTTAGAAAGCGCTCAACGAATTAACCCGCTGATTGTTTAAAGT 2881  
Db 5118 AATTGATGGCAGCGGTTAGAAAGCGCTCAACGAATTAACCCGCTGATTGTTTAAAGT 5177  
Qy 2882 GCTGGCGCTGTTCCGTTTATCAATCCATCAAGAAATTCGCAATTTCCGAGAGGAAGAGA 2941





Db 3604 CGGGTGAAGCGGGATTGCA-----ATTA 3627  
Qy 1142 AATAAAAAAGCACCCTGAAAAGGTGCTTTTTTTTGTGATGGTTTGAACCTGTTCTTTCTTA 1201  
Db 3628 CAATCACGCTGACACAGTAGTGGGTGATATGGATCATTTGCTGCMAACATTTTCATTCAA 3687  
Qy 1202 TCTTGATACATATAGAAATAACGTCATTTTATTTTATTTTATTTTATTTTATTTTGTCTGAAAGGTGCGT 1261  
Db 3688 TCGNAAGGAGCTGCACATCATTTGAGGGGTTTTTTTTTGTGACGGGAAGCTGA----- 3738  
Qy 1262 TGAAGTGTGGTATGATGTTTAAAGTATGAAAAACCTTAAATTTGGTTGACAGA 1321  
Db 3739 -----TACAGAAGTAGGGAAAAACGGTTATATCCAGCGGCTTGTG- 3776  
Qy 1322 AAAACCCCATCTGTTAAAGTTTATAAGTGACCTAAACAAATACTAAATAGATGGGGGTTTC 1381  
Db 3777 -----CTGCTTATTGAAAGACAATAATAGACATGTCGGGGGTGA 3816  
Qy 1382 TTTTAAATATTATGTCTCCTAATAGTAGCATTTTATTCAGATGAAAAAATCAAGGGTTTTAGT 1441  
Db 3817 TAAACCATTTTAAAGCGGGATATCGGCCATCATCCAGATAGTATACNAGTTTTGCTGAA 3876  
Qy 1442 GGACAAGACAAAAAGTGGAAAAAGTGAGACCATGTGCTTTAGGAAGACGAGTTTATTAATAGC 1501  
Db 3877 AGATATGTGCGAGA----- 3890  
Qy 1502 TGAATAAGAACGGTGCTCTCCAAATATTCCTTTATTTAGAAAGCAAACTTAAATTTACTG 1561  
Db 3891 -----CGAGCTTTTCATGAAGACATATCGCCTTTTGCTTCAA 3930  
Qy 1562 AAAAGGGAATGAGAATAGTAGAATGACCAATAATAATGACTAGAGAAAGAAAGATGAAGA 1621  
Db 3931 GCGCGCGCTTGCACCATACGTTGCAGGGAACCTTGAGGGAAGACTGTCACCATGGAAGA 3990  
Qy 1622 TTGTTCTGAATAATTAAGAAACGAATATTTGGATAAATATGGGGATGATGTTAAGGCTATTG 1681  
Db 3991 GGTTTAAAGCAATTTGGGGCGGATTTAGAGAAAA-----CATGAATGCTTCATCGTAG 4043  
Qy 1682 GTGTTTATGGCTCTCTGGTGTGAGCTGATGGCCCTATTTCGGATATTTCAGATGATGT 1741  
Db 4044 AAGGTGACGGCGGTATTTCTGTGCCATTTGGGAGAGGACTATTTGGTGAGTCAATGTCATAA 4103  
Qy 1742 GTGTCATGTCAACAGAGGAAGCAGAGTTCAGCCCATGAATGACAAACCGGTGAGTGGGAAG 1801  
Db 4104 -----AAGCGTTGCAGCTTCCCATGATTATTTGGGCGCTCTCGCCTTGGAAACC 4153  
Qy 1802 TGAAGTGAATTTTGTATAGGAAGAGATTTACTAGATTTATGATCTCAGGTGGAATCAG 1861  
Db 4154 ATTAATCATACCTTTTAACTGTCAAATATATGCAGAAAGCATGGGGCTTCCCAATCGCCGGA 4213  
Qy 1862 ATTGGCCGCTTACATGGTCAATTTTCTCTATTTTTCCGATTTTATGATTCAGGTGGAT 1921  
Db 4214 ATTATCATCAATGGAATCAGTGACTCTCCTGATGAAGATGAATAAAGCAATCTCTGAGATG 4273  
Qy 1922 ACTTAGAGAAAGTGTATCAAACTGCTMAATCGGTAGAACCCCAACGTTCCACAGATGCGA 1981  
Db 4274 ATTGAGCGCTTATGGGGTGCAGATTTTAGGGGTACGCCAAGACTTGCACACGTG--- 4330  
Qy 1982 TTTGTGCCCTTATCGTAGAAGAGCTGTTTGAATATGAGGCAAAATGCGGTAAATATTCGTG 2041  
Db 4331 -----ACGAAGAAACCGGTTCTACATATGTTAAAGACCATA----- 4367  
Qy 2042 TGCAAGGACGACACATCTTACCATCTTGTACGTGACAGGTAGCATGCGAGGTGCCA 2101  
Db 4368 -----TCAATCTATCATTTACTGATGAATCAAGTGGGGGTATGA 4405  
Qy 2102 TGTTGATTGGTCTGCATCATCGCATCTGTTTATAGCAGAGCGCTTCGGTCTTTAACTGAAG 2161  
Db 4406 GAATGAATCAATGATGGAATCTCGACACCGGTTGCTGGCTGGAGCAGAGTGAATGACG 4465  
Qy 2162 CAGTTAAGCAATCAGATCTTCTTCAGGTTAAGACCAATCTGTGTCAGTTGCTGAATGCTG 2221  
Db 4466 AAGAGGGCGCTTTCAATATTAATTTACTGCTGATGAAGATATTTTGTCTATTAATGACCGGG 4525

Qy 2222 GTCAACCTTTCCGACTCTGAGAAACCTTCTGGATCGCTAGAGAAATTTCTGGAATGGGATTC 2281  
Db 4526 CTTTTTCATCATAGAAAAACCTTTTTCGAAAAAAAAGTAAGTCAATATGATTTATGATG 4585  
Qy 2282 AGGAGTGGACAGAACGACACGAGTATATAGTGGATGTGTCAAAAACGATACCATATTTTGA 2341  
Db 4586 CGAAATCCGGGCTCTGCCCGGAAAACTGCGGCTATTGTTTCA-----GTCTGCGAT 4637  
Qy 2342 TTGAAAAGCGCGGATTTGAGTCTTTACCGGATGGTGAATAAGAAAAACGCTGCTTGAAGGCGC 2401  
Db 4638 TTCGAAAGCGCGGATTTGAGTCTTTACCGGATGGTGAATAAGAAAAACGCTGCTTGAAGGCGC 4697  
Qy 2402 GAAGCGGCGCACGATCTGAAATATCGGCACATATTTGTATCGTGGCAAGCGCAGAGGTCC 2461  
Db 4698 GAAGCGGCGCACGATCTGAAATATCGGCACATATTTGTATCGTGGCAAGCGCAGAGGTCC 4757  
Qy 2462 GTCTAAACAGAGAAGTGGATCAGGTCTGATAGTCGGTTTTCAGGAAATTAAGAGAGAGTATGG 2521  
Db 4758 GTCTAAACAGAGAAGTGGATCAGGTCTGATAGTCGGTTTTCAGGAAATTAAGAGAGAGTATGG 4817  
Qy 2522 ACTGAAGATTTGTGATGCTTGTGACTGTTGAAGCCAGAGCAGCGGAGAGCGGCTCAAAGA 2581  
Db 4818 ACTGAAGATTTGTGATGCTTGTGACTGTTGAAGCCAGAGCAGCGGAGAGCGGCTCAAAGA 4877  
Qy 2582 TGCAAGAGTAGACCGCTATATATTAATTTGAATACGTCAAGAGAAACCATTTCAAACAT 2641  
Db 4878 TGCAGGAGTAGACCGCTATATTAATTTGAATACGTCAAGAGAAACCATTTCAAACAT 4937  
Qy 2642 CACAACCTCATACTACATACGATGACAGTCAATACCGTTTGAATTCGCAAAAGAAATCGGG 2701  
Db 4938 CACAACCTCATACTACATACGATGACAGTCAATACCGTTTGAATTCGCAAAAGAAATCGGG 4997  
Qy 2702 GCTGTCTCGTGTTCAGGCGCCATTTATCGGATGAAGAGACGAAACAGGATGTCATTGA 2761  
Db 4998 GCTGTCTCGTGTTCAGGCGCCATTTATCGGATGAAGAGACGAAACAGGATGTCATTGA 5057  
Qy 2762 CATCGCCAAAGCTTTGAAGGCTCTTGACGCGGATTCATTCCTGTGAATTTTTTGTGATGC 2821  
Db 5058 CATCGCCAAAGCTTTGAAGGCTCTTGACGCGGATTCATTCCTGTGAATTTTTTGTGATGC 5117  
Qy 2822 AATTGATGGCAGCGCTTAGAAGCGTCAACGAATTTAAACCCGCTGATTTGTTTAAAGT 2881  
Db 5118 AATTGATGGCAGCGCTTAGAAGCGTCAACGAATTTAAACCCGCTGATTTGTTTAAAGT 5177  
Qy 2882 GCTGGCGCTGTCGGTTTTATCAATCCATCAAAAGAAATTTCCGATTTCCGGAGAGAGA 2941  
Db 5178 GCTGGCGCTGTCGGTTTTATCAATCCATCAAAAGAAATTTCCGATTTCCGGAGAGAGA 5237  
Qy 2942 GGTCAATCTCCGACATTTGACGCCATTTAGGCGCTTTACGCGCAAACTCCATTTTGTGCG 3001  
Db 5238 GGTCAATCTCCGACATTTGACGCCATTTAGGCGCTTTACGCGCAAACTCCATTTTGTGCG 5297  
Qy 3002 AGACTACTTAACAACTGCGCGGCAAGAGAGACGAGGATCATATAAATGCTGAGTGAATTT 3061  
Db 5298 AGACTACTTAACAACTGCGCGGCAAGAGAGACGAGGATCATATAAATGCTGAGTGAATTT 5357  
Qy 3062 AGGCTTTGAAGTTGAATCAGTCGAAAGAAATGAAGGCTAGTTTAAAGTCGAAAAAGCTGAAA 3121  
Db 5358 AGGCTTTGAAGTTGAATCAGTCGAAAGAAATGAAGGCTAGTTTAAAGTCGAAAAAGCTGAAA 5417  
Qy 3122 GAATCAATAAAGCAATCGGTATGATGTCGAAT 3154  
Db 5418 GAATCAATAAAGCAATCGGTATGATGTCGAAT 5450

## RESULT 4

US-09-407-728-1

; Sequence 1, Application US/09407728

; Patent No. 6841366

; GENERAL INFORMATION:

; APPLICANT: Bower, Stanley Grant

; APPLICANT: Perkins, John B.

APPLICANT: Yocum, R. Rogers  
APPLICANT: Pero, Janice G.  
TITLE OF INVENTION: BIOTIN BIOSYNTHESIS IN BACILLUS  
TITLE OF INVENTION: SUBTILIS  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA: US/09/407,728  
APPLICATION NUMBER: US/09/407,728  
FILING DATE:

CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/239,430  
FILING DATE: May 6, 1994  
APPLICATION NUMBER: 08/084,709  
FILING DATE: June 25, 1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Freeman, John W.  
REGISTRATION NUMBER: 29,066  
REFERENCE/DOCKET NUMBER: 04599/004001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 8478  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
US-09-407-728-1

Query Match 51.1%; Score 1612.6; DB 3; Length 8478;  
Best Local Similarity 73.0%; Pred. No. 0;  
Matches 2303; Conservative 0; Mismatches 624; Indels 226; Gaps 9;

Qy	2	GATCCAGAGGTTACGAGCCTTGAAGATTGATTCCTGTTAAACGAGCGGTTACACAGAA	61
Db	2524	GATCCAGAGGTTACGAGCCTTGAAGATTGATTCCTGTTAAACGAGCGGTTACACAGAA	2593
Qy	62	TGAAGAAGCGCGGTACATCGTAACCTGCGGTCAATGGATGGAGCGCGGTTCCAGAGA	121
Db	2584	TGAAGAAGCGCGGTACATCGTAACCTGCGGTCAATGGATGGAGCGCGGTTCCAGAGA	2643
Qy	122	GGAATATGTATGGCGAAATCAACCGGTCTGGTCTCTCAACAAATATTTAGGGCTCGCAA	181
Db	2644	GGAATATGTATGGCGAAATCAACCGGTCTGGTCTCTCAACAAATATTTAGGGCTCGCAA	2703
Qy	182	GCGATAGCGTTTCATCGATCGACGCCCAACAGCAATTCGACCAATTTGGGACGAGGCA	241
Db	2704	GCGATAGCGTTTCATCGATCGACGCCCAACAGCAATTCGACCAATTTGGGACGAGGCA	2763
Qy	242	GCGGTTCAAGCTTTAAACGACAGCGCAATTCGGTCTGGCATGAAAGCTAGAAAAGAGATTG	301
Db	2764	GCGGTTCAAGCTTTAAACGACAGCGCAATTCGGTCTGGCATGAAAGCTAGAAAAGAGATTG	2823
Qy	302	CCAGCTTTAAACTGACAGAGCGCCCTGCTGTTTTCGAGCGGTACTTGGCCCAATGTCG	361
Db	2824	CCAGCTTTAAACTGACAGAGCGCCCTGCTGTTTTCGAGCGGTACTTGGCCCAATGTCG	2883
Qy	362	GTGTCCTTTCACTCTTCGCAGAAAAGGAGATGTCATTTAAGTGACCGCTCAATCATG	421
Db	2884	GTGTCCTTTCACTCTTCGCAGAAAAGGAGATGTCATTTAAGTGACCGCTCAATCATG	2943

Qy	422	CAAGTATGATCGAGCGGCTGCCGACTTTCTAAGCGTGATACAGTTGTTTATCGCATATTG	481
Db	2944	CAAGTATGATCGAGCGGCTGCCGACTTTCTAAGCGTGATACAGTTGTTTATCGCATATTG	3003
Qy	482	ATATGAATGATCTTGAACCAAGCTGAATGAACACACAGCGTTATCAGCGCGGTTTATCG	541
Db	3004	ATATGAATGATCTTGAACCAAGCTGAATGAACACACAGCGTTATCAGCGCGGTTTATCG	3063
Qy	542	TACAGACGAGTATTCAGCATGGATGCGCAATTCGCCCTCTTGTAGTCAGATCATCTCAC	601
Db	3064	TACAGACGAGTATTCAGCATGGATGCGCAATTCGCCCTCTTGTAGTCAGATCATCTCAC	3123
Qy	602	TTGCGAAACGCTATCATGCTTCTGTCGTTGATGATGCCACGCAACAGAGATTG	661
Db	3124	TTGCGAAACGCTATCATGCTTCTGTCGTTGATGATGCCACGCAACAGAGATTG	3183
Qy	662	GCGATTGCGGACAGGAAAGGATGAACTTCTGTTGGTGTGTTGTCGCGACATTTGTTATCGCA	721
Db	3184	GCGATTGCGGACAGGAAAGGATGAACTTCTGTTGGTGTGTTGTCGCGACATTTGTTATCGCA	3243
Qy	722	CCTTAAGCAAGCTGTTGGCGCGGAGAGGTTTGGCGGAGGATCAGCGGTCTTCATCG	781
Db	3244	CCTTAAGCAAGCTGTTGGCGCGGAGAGGTTTGGCGGAGGATCAGCGGTCTTCATCG	3303
Qy	782	ACTTTTGTCTGAACCATGCCAGAACATTTATCTTTCAAAACCGCTATTCGCGCAGCCAGCT	841
Db	3304	ACTTTTGTCTGAACCATGCCAGAACATTTATCTTTCAAAACCGCTATTCGCGCAGCCAGCT	3363
Qy	842	GTGCGGCTGCTCAGAGGCTTTCAACATCATTTGAGCGCAGGAGGAAAACGACAGCTTT	901
Db	3364	GTGCGGCTGCTCAGAGGCTTTCAACATCATTTGAGCGCAGGAGGAAAACGACAGCTTT	3423
Qy	902	TATTTTCTTATATCAGCATGATCAGAACCACTCTGGAAGATATGGGTATCTGTTGAAAG	961
Db	3424	TATTTTCTTATATCAGCATGATCAGAACCACTCTGGAAGATATGGGTATCTGTTGAAAG	3483
Qy	962	GAGATCACACACCGGATTTCTGTAGTCAATGGCGATGCCCATATAAACCGTCTTATTTG	1021
Db	3484	GAGATCACACACCGGATTTCTGTAGTCAATGGCGATGCCCATATAAACCGTCTTATTTG	3543
Qy	1022	CTGAAAACGTCAGGCGAGGAAATTTATGCTCTGCGCATTCGCGCGCAACCGTGGCG	1081
Db	3544	CTGAAAACGTCAGGCGAGGAAATTTATGCTCTGCGCATTCGCGCGCAACCGTGGCG	3603
Qy	1082	CGGTGAAAGCGGATTCGAAAGCTTTGGGCGAGCGTCCGATCAGGGAATGAGTTTATAA	1141
Db	3604	CGGTGAAAGCGGATTCGAAAGCTTTGGGCGAGCGTCCGATCAGGGAATGAGTTTATAA	3627
Qy	1142	AATAAAAAAGCACCTGAAAAGGTGCTTTTTTTTGTATGGTTTGAACCTGTTCTTTCTTA	1201
Db	3628	CAATCACGTCGTGACCAACAGATATGGTGATATTTGATCATTTGCTGCAAAACATTTTCA	3687
Qy	1202	TCTTGATACATATAGAAATAACGTCATTTTTTTTATTTTTTTTATTTTGTCTGTAAGGTCG	1261
Db	3688	TCGGAAGAGGAGCTGCACATCATTTTGGGGGTTTTTTTGTGACGCGGAACTGA-----	3738
Qy	1262	TGAAGTGTGTATGTATGTTTAAAGTATTGAAAACCTTTAAATTTGGTTGTCACAGA	1321
Db	3739	-----TACAGAGTAGGAAAACCGTTATATATCCAGCGTCTTG- 3776	
Qy	1322	AAAAACCCATCTGTTAAAGTTTATAAGTGACTTAAACAAATAAATAAGTAGAGGGGTTTC	1381
Db	3777	-----CTGCCCTTATTGAAAAGACAATAATAGACATGTCGCGGGTGTA 3816	
Qy	1382	TTTTTATATTTATGTCTCTTAATAGTAGCATTTTATCAGATGAAAAAATCAAGGTTTAGT	1441
Db	3817	TAAACCAATTTTAAAGCGGAGATATCGCGCCATCATCCAGATAGTATACAGTTTGTGAA	3876
Qy	1442	GGACAGACAAAAGTGGAAAAGTGAGACCATGTGCTTAGGAAGACGAGTTTATTAATAGC	1501
Db	3877	AGATATGTCGAGA----- 3890	

Qy 1502 TGAATAAGACGGTCTCTCCAAATATTCTTAATTTAGAAAAACAAATCTAAATATTATCTG 1561  
Db 3891 -----CCAGTCTTTTCATGAGACATTACGCCCTTTTGCTTCAA 3930  
Qy 1562 AAAAGGGAAATGAGATAGTGAATGGACCAATAATAATGACTAGAGAAAGAAATGAAGA 1621  
Db 3931 GCGCGCGCTTGACCAATACGTTGACGGGAAACTTTGAGGAAAGACTGTCACCATGGAAGA 3990  
Qy 1622 TTGTTTCGAATTAAGGAACGAATATTGGGATAAATATGGGGATGATGTTAAGGCTATTG 1681  
Db 3991 GGTTTTAAGCCATTGGGGCGGATTTAGAGAAAA-----CATGAATGCTTCATCGTAG 4043  
Qy 1682 GTGTTTATGGCTCTCTCGTGGTCAAGTGAATGGGCCCTATTTCGGATATTGAGATGATGT 1741  
Db 4044 AAGGTGACGGCGGTATTCTGTGCCATTGGGAGAGGACTATTGTCAGTCATGTCTATAA 4103  
Qy 1742 GTGTCATGTCAACAGAGGAAGAGATTCAGCCATGAATGACCAACCGGTGAGTGGGAAG 1801  
Db 4104 -----AAGCGTTGCAGCTTCCCATGATTATTGTGGCGCGCTCTCGGCTTTGGAACC 4153  
Qy 1802 TGGAAAGTGAATTTGATAGCAAGAGATTCTACTAGATTATGATCTCAGGTGGATCAG 1861  
Db 4154 ATTAATCATACCTTTTAACTGTCAAAATATGCAAGAAAGCATGGGGCTTCCAAATCGCCGGA 4213  
Qy 1862 ATTGGCCGCTTACACATGGTCAAATTTTCTCTATTTTTGGCCGATTTATGATTCAGGTGGAT 1921  
Db 4214 ATTATCATATGGAATCAGTGACTCTCTCTGATGAAGTGAAGAAACCAATCTTGAGATG 4273  
Qy 1922 ACTTAGAGAAAGTGTATCAAACTGCTAAATCGGTAGAACCCAAACGTTCCACGATGCGA 1981  
Db 4274 ATTGAGCGCTTATGCGGTGTCGATTTTAGGGGTTACGCCAAAGCTTGCCCAACGCTG--- 4330  
Qy 1982 TTTGTGCCCTTATCGTAGAAGAGCTGTTTGAATATGACGGCAAAATGGGTAATATTTCGTG 2041  
Db 4331 -----ACGAAGAAACCGGTTCTACATATGTTAAAGACCATTA----- 4367  
Qy 2042 TGCAAGGACCGACAACATTTCTACCATCTTTGACTGTACAGGTAGCAATGCGAGGTGCCA 2101  
Db 4368 -----TCAATCTATCATTTACTGATGAATCAAGTGGGGTATGA 4405  
Qy 2102 TGTGTAATGTCGATCATCGCATCTGTTTATACGACGAGCGCTTCGGTCTTAACTGAAG 2161  
Db 4406 GAATGAATCAATGATGTAATCGCAGACCGGGTCTGGCTGGAGCAGAAAGTCACTGACG 4465  
Qy 2162 CAGTTAGCAATCAGATCTTCTCAGGTTATGACCATCTGTCGACGTTCCGTAATGCTG 2221  
Db 4466 AAGAGCGCTTTCAATATTAATTGTCCTGATGAAGATATTGCTATTAAATGCAACGGGG 4525  
Qy 2222 GTCAACTTTCCGACTCTGAGAAACTTCTGGAAATCGCTAGAGAAATTTCTGGAAATGGGATTC 2281  
Db 4526 CTTTTCACATCAGAAACACTTTTACGGAAAAAAGTAAAGCTCAATATGATTATGAATG 4585  
Qy 2282 AGGAGTGACAGAACGACACGGATATATAGTGGATGTGTCAAAACGCAATACCATTTTGAA 2341  
Db 4586 CGAAATCGGGCTCTGCCCGGAAAACTCGGCTATTGTTTACA-----GTCTGGCAT 4637  
Qy 2342 TTCCAAAGCGCGGATGATCTTACCGGATGGTGAATGAAGAAACGCTGCTTGAAGCGC 2401  
Db 4638 TTCGAAAGCGCGGATGATCTTACCGGATGGTGAATGAAGAAACGCTGCTTGAAGCGC 4697  
Qy 2402 GAAGCGGCGCACGATCTGAATATCGGCACATATTGTATCTGTGCAAGCGCACAGGTGCC 2461  
Db 4698 GAAGCGGCGCACGATCTGAATATCGGCACATATTGTATCTGTGCAAGCGCACAGGTGCC 4757  
Qy 2462 GTCTAAACAGAGAGTGGATCAGGTGATAGTCCGGTTACAGAAATTAAGAGACGATATGG 2521  
Db 4758 GTCTAAACAGAGAGTGGATCAGGTGATAGTCCGGTTACAGAAATTAAGAGACGATATGG 4817  
Qy 2522 ACTGAAGATTTGCTGATGCTTGGACTGTTGAAGCCAGAGCAGCGGCTCAAGA 2581  
Db 4818 ACTGAAGATTTGCTGATGCTTGGACTGTTGAAGCCAGAGCAGCGGCTCAAGA 4877  
Qy 2582 TGCAGGAGTAGACCGCTATATATCATTAATTTGAATACGTCACAGAGAAACCATTTCAAACAT 2641

Db 4878 TGCAGGAGTAGACCGCTATAATCATTAATTTGAATACGTCACAGAGAAACCATTTCAAACAT 4937  
Qy 2642 CACAACCTCATACATACATACGATGACAGAGTCAATACACGTTGAAATCGCAAAAGAAATCGGG 2701  
Db 4938 CACAACCTCATACATACATACGATGACAGAGTCAATACACGTTGAAATCGCAAAAGAAATCGGG 4997  
Qy 2702 GCTGTCTCCGTGTTTCAGCGCCCATTTATCGGGATGAAGAGACGAAACAGGATGTCATTGA 2761  
Db 4998 GCTGTCTCCGTGTTTCAGCGCCCATTTATCGGGATGAAGAGACGAAACAGGATGTCATTGA 5057  
Qy 2762 CATGCCCAAAGCTTGAAGGCTCTTGACGCGGATTCATCTCTGTGAATTTTTTGCATGC 2821  
Db 5058 CATGCCCAAAGCTTGAAGGCTCTTGACGCGGATTCATCTCTGTGAATTTTTTGCATGC 5117  
Qy 2822 AATTGATGGCAGCGGTTAGAGCGCTCAACGAATTAACCCGCTGTATTGTTTAAAGT 2881  
Db 5118 AATTGATGGCAGCGGTTAGAGCGCTCAACGAATTAACCCGCTGTATTGTTTAAAGT 5177  
Qy 2882 GCTGGCGCTGTTCCGTTTTATCAATCCATCAAAAGAAATTCGCATTTCCGAGGAAGAGA 2941  
Db 5178 GCTGGCGCTGTTCCGTTTTATCAATCCATCAAAAGAAATTCGCATTTCCGAGGAAGAGA 5237  
Qy 2942 GGTCAATCTCCGCACATTGACGCCATTAGGCGCTTTACGCCGCAAACTCCATTTTGTGCG 3001  
Db 5238 GGTCAATCTCCGCACATTGACGCCATTAGGCGCTTTACGCCGCAAACTCCATTTTGTGCG 5297  
Qy 3002 AGACTACTTAACTACCTCCGGGACGAGGAGACGAGGATCATAAATGCTGAGTATTT 3061  
Db 5298 AGACTACTTAACTACCTCCGGGACGAGGAGACGAGGATCATAAATGCTGAGTATTT 5357  
Qy 3062 AGGCTTTGAAGTTGAATCAGTCGAAAGAAATGAAGGCTAGTTTAAGTCGAAAACTGAAA 3121  
Db 5358 AGGCTTTGAAGTTGAATCAGTCGAAAGAAATGAAGGCTAGTTTAAGTCGAAAACTGAAA 5417  
Qy 3122 GAATCAATAAAGCAATCGGTATGATGTCGAAT 3154  
Db 5418 GAATCAATAAAGCAATCGGTATGATGTCGAAT 5450

RESULT 5

US-09-869-855A-3  
; Sequence 3, Application US/09869855A  
; Patent No. 6762040  
; GENERAL INFORMATION:  
; APPLICANT: Rasmussen, Michael Dolberg  
; TITLE OF INVENTION: Method For Increasing Gene Copy Number  
; FILE REFERENCE: 10028.204-US  
; CURRENT APPLICATION NUMBER: US/09/869,855A  
; CURRENT FILING DATE: 2001-07-05  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 5793  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic  
; US-09-869-855A-3

Query Match 27.5%; Score 867; DB 3; Length 5793;  
Best Local Similarity 100.0%; Pred. No. 2.3e-237;  
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1475 TGCTTAGGAAGACGAGTTATTAATAGCTGAATAAGAACGGTGCTCTCCAAATATTTCTTAT 1534  
Db 968 TGCTTAGGAAGACGAGTTATTAATAGCTGAATAAGAACGGTGCTCTCCAAATATTTCTTAT 1027  
Qy 1535 TTAGAAAAGCAATCTAAATATATCTGAAAAGGAATGAGATAGTGAATGCAACCAATAA 1594  
Db 1028 TTAGAAAAGCAATCTAAATATATCTGAAAAGGAATGAGATAGTGAATGCAACCAATAA 1087  
Qy 1595 TAATGACTAGAGAAGAAAGAAATGAAGATTGTTTCATGAAATTTAAGGAACGAATATTGGATA 1654

|||||  
Db 1088 TAATGACTAGAGAGAAAGAAATGAAGATGTTTCATGAAATTAAGGAAACGAATATTGGATA 1147  
Qy 1655 AATATGGGATGATGTTAAAGCTATTTCGTTGTTATGCTCTCTTGGTCTGTCAGACTCATG 1714  
Db 1148 AATATGGGATGATGTTAAAGCTATTTCGTTGTTATGCTCTCTTGGTCTGTCAGACTCATG 1207  
Qy 1715 GGCCCTATTCCGATATTGAGATGATGTGTGTCATGTCAACAGAGGAAAGCAGAGTTTCAGCC 1774  
Db 1208 GGCCCTATTCCGATATTGAGATGATGTGTGTCATGTCAACAGAGGAAAGCAGAGTTTCAGCC 1267  
Qy 1775 ATGATGAGCAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTATAGCAAGAGATTCTTAC 1834  
Db 1268 ATGATGAGCAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTATAGCAAGAGATTCTTAC 1327  
Qy 1835 TAGATATTGATCTCAGTCCAGTGGATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 1894  
Db 1328 TAGATATTGATCTCAGTCCAGTGGATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 1387  
Qy 1895 TTTTGGCGATTTATGATTTCAAGTGGATACCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 1954  
Db 1388 TTTTGGCGATTTATGATTTCAAGTGGATACCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 1447  
Qy 1955 TAGAAGCCCAACGTTTCCAGATCGGATTTTGGCCCTTATCGTAGAAGAGCTGTTGAAT 2014  
Db 1448 TAGAAGCCCAACGTTTCCAGATCGGATTTTGGCCCTTATCGTAGAAGAGCTGTTGAAT 1507  
Qy 2015 ATGCAGGCAAAATGCGGTAAATATTCGTGCAAGGACCGCAACATTTCTACCATCCTTGA 2074  
Db 1508 ATGCAGGCAAAATGCGGTAAATATTCGTGCAAGGACCGCAACATTTCTACCATCCTTGA 1567  
Qy 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTTGCTGTCATCGCATCTGTTATA 2134  
Db 1568 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTTGCTGTCATCGCATCTGTTATA 1627  
Qy 2135 CGACGAGCGTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTTATG 2194  
Db 1628 CGACGAGCGTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTTATG 1687  
Qy 2195 ACCATCTGTGCCAGTTGCTAATGTCGTGCAAGCAGTCTTCCGACTCTGAGAACTTCTGGAAT 2254  
Db 1688 ACCATCTGTGCCAGTTGCTAATGTCGTGCAAGCAGTCTTCCGACTCTGAGAACTTCTGGAAT 1747  
Qy 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGAGTGGACAGAACGACACGGATATATAGTGG 2314  
Db 1748 CGCTAGAGAAATTTCTGGAATGGGATTCAGAGTGGACAGAACGACACGGATATATAGTGG 1807  
Qy 2315 ATGTGTCAAAACGCATACCAATTTTGAA 2341  
Db 1808 ATGTGTCAAAACGCATACCAATTTTGAA 1834

RESULT 6  
US-09-869-855A-2  
; Sequence 2, Application US/09869855A  
; Patent No. 6762040  
; GENERAL INFORMATION:  
; APPLICANT: Rasmussen, Michael Dolberg  
; TITLE OF INVENTION: Method For Increasing Gene Copy Number  
; FILE REFERENCE: 10028.204-US  
; CURRENT APPLICATION NUMBER: US/09/869,855A  
; CURRENT FILING DATE: 2001-07-05  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 5943  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic  
US-09-869-855A-2

Query Match 27.5%; Score 867; DB 3; Length 5943;

Best Local Similarity 100.0%; Pred. No. 2,3e-237;  
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1475 TGCTTAGGAAGACGAGTTATTAAATAGCTGAATTAAGAACGCTGCTCTCCAATATTCTTAT 1534  
Db 968 TGCTTAGGAAGACGAGTTATTAAATAGCTGAATTAAGAACGCTGCTCTCCAATATTCTTAT 1027  
Qy 1535 TTAGAAAAGCAATCTAAAAATTATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA 1594  
Db 1028 TTAGAAAAGCAATCTAAAAATTATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA 1087  
Qy 1595 TAATGACTAGAGAAGAAAGAAATGAAGATTGTTTCATGAAATTAAGGAACGAATATTGGATA 1654  
Db 1088 TAATGACTAGAGAAGAAAGAAATGAAGATTGTTTCATGAAATTAAGGAACGAATATTGGATA 1147  
Qy 1655 AATATGGGATGATGTTAAAGCTATTTCGTTGTTATGCTCTCTTGGTCTGTCAGACTCATG 1714  
Db 1148 AATATGGGATGATGTTAAAGCTATTTCGTTGTTATGCTCTCTTGGTCTGTCAGACTCATG 1207  
Qy 1715 GGCCCTATTCCGATATTGAGATGATGTGTGTCATGTCAACAGAGGAAAGCAGAGTTTCAGCC 1774  
Db 1208 GGCCCTATTCCGATATTGAGATGATGTGTGTCATGTCAACAGAGGAAAGCAGAGTTTCAGCC 1267  
Qy 1775 ATGAATGGAACCAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTATAGCAAGAGATTCTTAC 1834  
Db 1268 ATGAATGGAACCAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTATAGCAAGAGATTCTTAC 1327  
Qy 1835 TAGATATTGATCTCAGTCCAGTGGATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 1894  
Db 1328 TAGATATTGATCTCAGTCCAGTGGATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA 1387  
Qy 1895 TTTTGGCGATTTATGATTTCAAGTGGATACCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 1954  
Db 1388 TTTTGGCGATTTATGATTTCAAGTGGATACCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 1447  
Qy 1955 TAGAAGCCCAACGTTTCCAGATCGGATTTTGGCCCTTATCGTAGAAGAGCTGTTGAAT 2014  
Db 1448 TAGAAGCCCAACGTTTCCAGATCGGATTTTGGCCCTTATCGTAGAAGAGCTGTTGAAT 1507  
Qy 2015 ATGCAGGCAAAATGCGGTAAATATTCGTGCAAGGACCGCAACATTTCTACCATCCTTGA 2074  
Db 1508 ATGCAGGCAAAATGCGGTAAATATTCGTGCAAGGACCGCAACATTTCTACCATCCTTGA 1567  
Qy 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTTGCTGTCATCGCATCTGTTATA 2134  
Db 1568 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTTGCTGTCATCGCATCTGTTATA 1627  
Qy 2135 CGACGAGCGTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTTATG 2194  
Db 1628 CGACGAGCGTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTTATG 1687  
Qy 2195 ACCATCTGTGCCAGTTGCTAATGTCGTGCAAGCAGTCTTCCGACTCTGAGAACTTCTGGAAT 2254  
Db 1688 ACCATCTGTGCCAGTTGCTAATGTCGTGCAAGCAGTCTTCCGACTCTGAGAACTTCTGGAAT 1747  
Qy 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGAGTGGACAGAACGACACGGATATATAGTGG 2314  
Db 1748 CGCTAGAGAAATTTCTGGAATGGGATTCAGAGTGGACAGAACGACACGGATATATAGTGG 1807  
Qy 2315 ATGTGTCAAAACGCATACCAATTTTGAA 2341  
Db 1808 ATGTGTCAAAACGCATACCAATTTTGAA 1834

RESULT 7  
US-08-418-085-3  
; Sequence 3, Application US/08418085  
; Patent No. 5869283  
; GENERAL INFORMATION:  
; APPLICANT: SLIJKHUIS, HERMAN; SELTEN, GERARDUS CORNELIS  
; APPLICANT: MARIA; SMAAL, ERIC BASTIAN  
; TITLE OF INVENTION: PROCESS FOR OXIDATION OF STEROIDS AND  
; TITLE OF INVENTION: GENETICALLY ENGINEERED CELLS USED THEREIN

NUMBER OF SEQUENCES: 79  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BIERMAN & MUSERLIAN  
STREET: 600 THIRD AVENUE  
CITY: NEW YORK  
STATE: NEW YORK  
COUNTRY: USA  
ZIP: 10016  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: IBM PC COMPATIBLE  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/418,085  
FILING DATE: 06-APR-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/054,185  
FILING DATE: 26-APR-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/002,608  
FILING DATE: 11-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/474,798  
FILING DATE: 16-JULY-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/NL89/00072  
FILING DATE: 25-SEPT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: NL/88/202080.3  
FILING DATE: 03-SEP-88  
ATTORNEY/AGENT INFORMATION:  
NAME: CHARLES A. MUSERLIAN  
REGISTRATION NUMBER: 19,683  
REFERENCE/DOCKET NUMBER: 146.1169 CON-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 661-8000  
TELEFAX: (212) 661-8002  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7336  
TYPE: NUCLEIC ACID  
STRANDEDNESS: DOUBLE  
TOPOLOGY: UNKNOWN  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: PLASMID pBHA-1  
US-08-418-085-3  
Query Match 27.5%; Score 867; DB 2; Length 7336;  
Best Local Similarity 100.0%; Pred. No. 2.6e-237;  
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1475 TGCTTAGGAACGAGTTATTAAATAGCTGAATAAGACGGTGCTCTCCAAATATCTTAT 1534  
DB 5231 TGCTTAGGAACGAGTTATTAAATAGCTGAATAAGACGGTGCTCTCCAAATATCTTAT 5290  
QY 1535 TTAGAAAGCAATCTAAATTTATCTGAAAGGGAATGAGAAATAGTGAATGACCAATAA 1594  
DB 5291 TTAGAAAGCAATCTAAATTTATCTGAAAGGGAATGAGAAATAGTGAATGACCAATAA 5350  
QY 1595 TAATGACTAGAGAAGAAAGATTGTTTCATGAAATTAAGGAACGAATATTGGATA 1654  
|||||

DB 5351 TAATGACTAGAGAAGAAAGATTGTTTCATGAAATTAAGGAACGAATATTGGATA 5410  
QY 1655 AATATGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCTTTGGTGGTCTGAGACTGATG 1714  
DB 5411 AATATGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCTTTGGTGGTCTGAGACTGATG 5470  
QY 1715 GGCCCTATTCCGATATTGAGATGATGTGTCTGATGTCATGTCACAGAGAGAAACAGAGTTTCAGCC 1774  
DB 5471 GGCCCTATTCCGATATTGAGATGATGTGTCTGATGTCACAGAGAGAAACAGAGTTTCAGCC 5530  
QY 1775 ATGAATGGAACAACCGGTGAGTGAAGGTGAAGTGAATTTTGTATAGGAGAAAGAGATTCTTAC 1834  
DB 5531 ATGAATGGAACAACCGGTGAGTGAAGGTGAAGTGAATTTTGTATAGGAGAAAGATTCTTAC 5590  
QY 1835 TAGATTATGCATCTCAGGTGGAAATCAGATTGGCCGCTTACATGTTGCTCAATTTTCTCTA 1894  
DB 5591 TAGATTATGCATCTCAGGTGGAAATCAGATTGGCCGCTTACATGTTGCTCAATTTTCTCTA 5650  
QY 1895 TTTTGGCCGATTTATGATTTCAGGTGGATCTTATAGAGAAAGTGTATCAAACTGCTAAATCGG 1954  
DB 5651 TTTTGGCCGATTTATGATTTCAGGTGGATCTTATAGAGAAAGTGTATCAAACTGCTAAATCGG 5710  
QY 1955 TAGAAGCCCAAAACGTTCCAGATGCGATTTGTGCCCTTATCGTAGAAGAGCTGTTTGAAT 2014  
DB 5711 TAGAAGCCCAAAACGTTCCAGATGCGATTTGTGCCCTTATCGTAGAAGAGCTGTTTGAAT 5770  
QY 2015 ATGCAGGCAAAATGGCGTAATATTTCGTGTGCAAGACCGCAACAATTTCTACCATCCTTGA 2074  
DB 5771 ATGCAGGCAAAATGGCGTAATATTTCGTGTGCAAGACCGCAACAATTTCTACCATCCTTGA 5830  
QY 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTTGCTGTGCATCATCGCATCTGTTATA 2134  
DB 5831 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTTGCTGTGCATCATCGCATCTGTTATA 5890  
QY 2135 CGACGACGCTTCGGTCTTAACTGAACAGTGAAGCAATCAGATCTTCCTTCAGGTTATG 2194  
DB 5891 CGACGACGCTTCGGTCTTAACTGAACAGTGAAGCAATCAGATCTTCCTTCAGGTTATG 5950  
QY 2195 ACCATCTGTCCAGTTTCGTAAATGTCGTCAACTTTCCGACTCTGAGAACTTCCTGGAAT 2254  
DB 5951 ACCATCTGTCCAGTTTCGTAAATGTCGTCAACTTTCCGACTCTGAGAACTTCCTGGAAT 6010  
QY 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGACGACACGGATATATAGTGG 2314  
DB 6011 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGACGACACGGATATATAGTGG 6070  
QY 2315 ATGTGTCAAAACGCATACCATTTTGA 2341  
DB 6071 ATGTGTCAAAACGCATACCATTTTGA 6097

## RESULT 8

US-09-099-011A-3  
; Sequence 3, Application US/09099011A  
; Patent No. 6171836  
; GENERAL INFORMATION:  
; APPLICANT: SLIJKHUIS, HERMAN; SELTEN,  
; APPLICANT: GERARDUS CORNELIS MARIA; SMAL,  
; APPLICANT: ERIC BASTIAAN  
; TITLE OF INVENTION: PROCESS FOR OXIDATION OF  
; TITLE OF INVENTION: STEROIDS AND GENETICALLY ENGINEERED CELLS  
; NUMBER OF INVENTION: USED THEREIN  
; NUMBER OF SEQUENCES: 79  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BIERMAN, MUSERLIAN & LUCAS  
; STREET: 600 THIRD AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10016  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MICROSOFT WORD 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/099,011A
; FILING DATE: 17-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/418,085
; FILING DATE: 06-APR-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/054,185
; FILING DATE: 26-APR-1993
; APPLICATION DATA:
; APPLICATION NUMBER: 08/002,608
; FILING DATE: 11-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/474,857
; FILING DATE: 30-OCT-1990
; APPLICATION DATA:
; APPLICATION NUMBER: 07/474,798
; FILING DATE: 16-JULY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/NL89/00072
; FILING DATE: 25-SEPT-1989
; APPLICATION DATA:
; APPLICATION NUMBER: NL88/200904.6
; FILING DATE: 06-MAY-1988
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: NL/88/202080.3
; FILING DATE: 03-SEP-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: CHARLES A. MUSERLIAN
; REGISTRATION NUMBER: 19,683
; REFERENCE/DOCKET NUMBER: 146.1169-
; REFERENCE/DOCKET NUMBER: CON-1-DIV-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 661-8000
; TELEFAX: (212) 661-8002
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7336 BASE PAIRS
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: UNKNOWN
; FEATURE:
; OTHER INFORMATION: PLASMID pBHA-1
; US-09-099-011A-3

Query Match 27.5%; Score 867; DB 3; Length 7336;
Best Local Similarity 100.0%; Pred. No. 2.6e-237;
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1475 TGCTTAGGACGAGTATTATAGCTGAATAAGACGGTGTCTTCCAAATATTTCTAT 1534
DB 5231 TGCTTAGGACGAGTATTATAGCTGAATAAGACGGTGTCTTCCAAATATTTCTAT 5290
QY 1535 TTAGAAAAGCAATCTAAAATTTCTGAAAAGGAATGAGATAGTGAATGGACCAATAA 1594
DB 5291 TTAGAAAAGCAATCTAAAATTTCTGAAAAGGAATGAGATAGTGAATGGACCAATAA 5350
QY 1595 TAATGACTAGAGAAGAAGATGAAGATTTTCATGAAATTAAGGAACGAATATTGGATA 1654
DB 5351 TAATGACTAGAGAAGAAGATGAAGATTTTCATGAAATTAAGGAACGAATATTGGATA 5410
QY 1655 AATATGGGATGATTTAAGGCTATTGGTCTTTATGCTCTCTTGGTCGTCAGACTGATG 1714
DB 5411 AATATGGGATGATTTAAGGCTATTGGTCTTTATGCTCTCTTGGTCGTCAGACTGATG 5470
QY 1715 GGCCCTATTCGGATATTGAGATGATGTGTCATGTCACAGAGGAGCAGAGTTTCAGCC 1774
DB 5471 GGCCCTATTCGGATATTGAGATGATGTGTCATGTCACAGAGGAGCAGAGTTTCAGCC 5530
QY 1775 ATGAATGGACAACCGGTGAGTGAAGGTGAAGTGAATTTTGTAGCGAAGAGATTCTAC 1834
;

5531 ATCAATGGACAACCGGTGAGTGAAGGTGAAGTGAATTTTGTAGCGAAGAGATTCTAC 5590
1835 TAGATTATGATCTCAGGTGGAATCAGATTGGCGGCTTACACATGGTCAATTTTCTCTA 1894
5591 TAGATTATGATCTCAGGTGGAATCAGATTGGCGGCTTACACATGGTCAATTTTCTCTA 5650
1895 TTTTGGCGATTTATGATTCAAGTGGATCTATAGAGAAAGTGTATCAAACTGCTAAATCGG 1954
5651 TTTTGGCGATTTATGATTCAAGTGGATCTATAGAGAAAGTGTATCAAACTGCTAAATCGG 5710
1955 TAGAAGCCCAACGTTCCACGATGCGATTGTGCGCCCTTATCGTAGAAGAGCTGTTGAAT 2014
5711 TAGAAGCCCAACGTTCCACGATGCGATTGTGCGCCCTTATCGTAGAAGAGCTGTTGAAT 5770
2015 ATCAGGCAAAATGCGGTAAATATTCTGTGTCGCAAGCCGACCAACATTTCTACCATCCTTGA 2074
5771 ATCAGGCAAAATGCGGTAAATATTCTGTGTCGCAAGCCGACCAACATTTCTACCATCCTTGA 5830
2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTTGCTGTCATCATCGCATCTGTTATA 2134
5831 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTTGCTGTCATCATCGCATCTGTTATA 5890
2135 CGACGAGCGTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTTATG 2194
5891 CGACGAGCGTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTTATG 5950
2195 ACCATCTGTGCCAGTTGCTGAATCTGTGGTCAACTTTCCGACTCTGAGAAAATCTTCGGAAT 2254
5951 ACCATCTGTGCCAGTTGCTGAATCTGTGGTCAACTTTCCGACTCTGAGAAAATCTTCGGAAT 6010
2255 CGCTAGAGAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACGAGATATATAGTGG 2314
6011 CGCTAGAGAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACGAGATATATAGTGG 6070
2315 ATGTGTCAAAAACGCATACCATTTTGA 2341
6071 ATGTGTCAAAAACGCATACCATTTTGA 6097

RESULT 9
US-08-470-369-21
; Sequence 21, Application US/08470369
; Patent No. 623889
; GENERAL INFORMATION:
; APPLICANT: Dorsers, Lambertus C. J.
; APPLICANT: Wagemaker, Gerard
; APPLICANT: Vos, Yvonne J.
; APPLICANT: Van Leen, Robert W.
; TITLE OF INVENTION: MOLECULAR CLONING AND EXPRESSION OF HUMAN
; TITLE OF INVENTION: IL-3
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/470,369
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/321,480
; FILING DATE:
; APPLICATION NUMBER: 07/854,297
; FILING DATE: 19-MAR-1992
; ATTORNEY/AGENT INFORMATION:

```





; TOPOLOGY: UNKNOWN

; FEATURE:

; NAME/KEY:

; LOCATION:

; IDENTIFICATION METHOD:

; OTHER INFORMATION: PLASMID pBHA-1

US-09-098-877B-3

Query Match 27.5%; Score 867; DB 3; Length 7336;  
Best Local Similarity 100.0%; Pred. No. 2.6e-237; Indels 0; Gaps 0;  
Matches 867; Conservative 0; Mismatches 0;

Qy 1475 TGCTTAGGAAGACAGATTATTAATAGCTGAATAAGACGGTCTCTCAAATATTTCTAT 1534  
Db 5231 TGCTTAGGAAGACAGATTATTAATAGCTGAATAAGACGGTCTCTCAAATATTTCTAT 5290

Qy 1535 TTAGAAAGCAATCTAAATATCTGAAAGGAATGAGATAGTGAATGGACCAATAA 1594  
Db 5291 TTAGAAAGCAATCTAAATATCTGAAAGGAATGAGATAGTGAATGGACCAATAA 5350

Qy 1595 TAATGACTAGAGAAGAAAGATGAAGATTGTTCAATGAAATTAAGGAACGAATATTGGATA 1654  
Db 5351 TAATGACTAGAGAAGAAAGATGAAGATTGTTCAATGAAATTAAGGAACGAATATTGGATA 5410

Qy 1655 AATATGGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCTTGGTCGTGAGACTGATG 1714  
Db 5411 AATATGGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCTTGGTCGTGAGACTGATG 5470

Qy 1715 GGCCCTATTGGATATTGAGATGATGTTGTCATGTCATCAACAGAGGAAGAGATTGAGCC 1774  
Db 5471 GGCCCTATTGGATATTGAGATGATGTTGTCATGTCATCAACAGAGGAAGAGATTGAGCC 5530

Qy 1775 ATGAATGCACAAACCGGTGAGTGAAGGTGGAAGTGAATTTTGATAGCGAAGAGATTCTAC 1834  
Db 5531 ATGAATGCACAAACCGGTGAGTGAAGGTGGAAGTGAATTTTGATAGCGAAGAGATTCTAC 5590

Qy 1835 TAGATATTGCTATCTCAGGTGGAATCAGATTGGCCGCTTACACATGTTCAATTTTCTCTA 1894  
Db 5591 TAGATATTGCTATCTCAGGTGGAATCAGATTGGCCGCTTACACATGTTCAATTTTCTCTA 5650

Qy 1895 TTTTGCCGATTTATGATTCAGGTGGATCACTTAGAGAAAGTGTATCAAACTGCTAAATCGG 1954  
Db 5651 TTTTGCCGATTTATGATTCAGGTGGATCACTTAGAGAAAGTGTATCAAACTGCTAAATCGG 5710

Qy 1955 TAGAAGCCAAACCGTTCACGATCGGATTTGTCGCCCTTATCGTAGAAGAGCTGTTTGAAT 2014  
Db 5711 TAGAAGCCAAACCGTTCACGATCGGATTTGTCGCCCTTATCGTAGAAGAGCTGTTTGAAT 5770

Qy 2015 ATGAGGCAAAATGGCGTAATATTCGTGTGCAAGGACCGACAACTTTCTACCATCTTGA 2074  
Db 5771 ATGAGGCAAAATGGCGTAATATTCGTGTGCAAGGACCGACAACTTTCTACCATCTTGA 5830

Qy 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTTGTTGTCATCATCGCATCTGTTATA 2134  
Db 5831 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTTGTTGTCATCATCGCATCTGTTATA 5890

Qy 2135 CGACGAGCGTTCGGTCTTAACTGAACGATTAAGCAATCAGATCTTCTTCCGTTTATG 2194  
Db 5891 CGACGAGCGTTCGGTCTTAACTGAACGATTAAGCAATCAGATCTTCTTCCGTTTATG 5950

Qy 2195 ACCATCTGTGCGAGTTCGTAAATGTCGTCTCAACTTTTCCGACTCTGAGAACTTCTGGAAT 2254  
Db 5951 ACCATCTGTGCGAGTTCGTAAATGTCGTCTCAACTTTTCCGACTCTGAGAACTTCTGGAAT 6010

Qy 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAAACGACAGGATATATAGTGG 2314  
Db 6011 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAAACGACAGGATATATAGTGG 6070

Qy 2315 ATGTGTCAAACGCAATACCATTTTGA 2341  
Db 6071 ATGTGTCAAACGCAATACCATTTTGA 6097

RESULT 11

US-08-460-343B-1

; Sequence 1, Application US/08460343B

; Patent No. 5741664

; GENERAL INFORMATION:

; APPLICANT: Marcus D. Ballinger and James A. Wells

; TITLE OF INVENTION: SUBSTITUTED VARIANTS CAPABLE OF CLEAVING

; TITLE OF INVENTION: SUBSTRATES CONTAINING DIBASIC RESIDUES

; NUMBER OF SEQUENCES: 74

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 1 DNA Way

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WinPatIn (Genentech)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/460,343B

; FILING DATE: 01-Jun-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/398028

; FILING DATE: 03-mar-1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Kubinec, Jeffrey S.

; REGISTRATION NUMBER: 36,575

; REFERENCE/DOCKET NUMBER: P0936C1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650/225-8228

; TELEFAX: 650/952-9881

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 8119 base pairs

; TYPE: Nucleic Acid

; STRANDEDNESS: Single

; TOPOLOGY: Linear

US-08-460-343B-1

Query Match 27.5%; Score 867; DB 2; Length 8119;

Best Local Similarity 100.0%; Pred. No. 2.8e-237;

Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1475 TGCTTAGGAAGACAGATTATTAATAGCTGAATAAGACGGTCTCTCAAATATTTCTAT 1534  
Db 5764 TGCTTAGGAAGACAGATTATTAATAGCTGAATAAGACGGTCTCTCAAATATTTCTAT 5823

Qy 1535 TTAGAAAGCAATCTAAATATCTGAAAGGAATGAGATAGTGAATGGACCAATAA 1594  
Db 5824 TTAGAAAGCAATCTAAATATCTGAAAGGAATGAGATAGTGAATGGACCAATAA 5883

Qy 1595 TAATGACTAGAGAAGAAAGATGAAGATTGTTCAATGAAATTAAGGAACGAATATTGGATA 1654  
Db 5884 TAATGACTAGAGAAGAAAGATGAAGATTGTTCAATGAAATTAAGGAACGAATATTGGATA 5943

Qy 1655 AATATGGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCTTGGTCGTGAGACTGATG 1714  
Db 5944 AATATGGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCTTGGTCGTGAGACTGATG 6003

Qy 1715 GGCCCTATTGGATATTGAGATGATGTCATGTCATCAACAGAGGAAGAGATTGAGCC 1774  
Db 6004 GGCCCTATTGGATATTGAGATGATGTCATGTCATCAACAGAGGAAGAGATTGAGCC 6063

Qy 1775 ATGAATGGACAAACCGGTGAGTGAAGGTGGAAGTGAATTTTGATAGCGAAGAGATTCTAC 1834  
Db 6064 ATGAATGGACAAACCGGTGAGTGAAGGTGGAAGTGAATTTTGATAGCGAAGAGATTCTAC 6123

Qy 1835 TAGATATTGCTATCTCAGGTGGAATCAGATTGGCCGCTTACATGTTGTTTCTCTA 1894

Db 6124 TAGATTATGCTATCTCAGGTGGAATCAGATTGGCCGCTTTACACATGGTCAATTTTCTCTA 6183  
Qy 1895 TTTTGGCGATTATGATTCAAGTGGATCTTATAGAGAAAGTGTATCAAACTCTCTAAATCGG 1954  
Db 6184 TTTTGGCGATTATGATTCAAGTGGATCTTATAGAGAAAGTGTATCAAACTCTCTAAATCGG 6243  
Qy 1955 TAGAAGCCCAACCGTTCCACGATGCGATTTCGTGCGCTTATCGTAGAGAGCTGTTTGAAT 2014  
Db 6244 TAGAAGCCCAACCGTTCCACGATGCGATTTCGTGCGCTTATCGTAGAGAGCTGTTTGAAT 6303  
Qy 2015 ATGCAGGCAAAATGCGGTAAATTCGTGCGAAGACCGACAACTTTCTACCATCTCTTGA 2074  
Db 6304 ATGCAGGCAAAATGCGGTAAATTCGTGCGAAGACCGACAACTTTCTACCATCTCTTGA 6363  
Qy 2075 CTGTACAGGTAGCAATGCGAGTGGCCATGTTGATTGGTCTGCATCATCGCATCTGTTATA 2134  
Db 6364 CTGTACAGGTAGCAATGCGAGTGGCCATGTTGATTGGTCTGCATCATCGCATCTGTTATA 6423  
Qy 2135 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTTATG 2194  
Db 6424 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTTATG 6483  
Qy 2195 ACCATCTGTGCCAGTTCGTAATGTCGTCAACTTTCCGACTCTGAGAACTTCTGGAAT 2254  
Db 6484 ACCATCTGTGCCAGTTCGTAATGTCGTCAACTTTCCGACTCTGAGAACTTCTGGAAT 6543  
Qy 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGACCGATATATAGTGG 2314  
Db 6544 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGACCGATATATAGTGG 6603  
Qy 2315 ATGTGTCAAAACGCATACCATTTTGAA 2341  
Db 6604 ATGTGTCAAAACGCATACCATTTTGAA 6630

## RESULT 12

US-08-398-028B-1

; Sequence 1, Application US/08398028B

; Patent No. 5780285

; GENERAL INFORMATION:

; APPLICANT: Marcus D. Ballinger and James A. Wells

; TITLE OF INVENTION: SUBSTITUTED VARIANTS CAPABLE OF CLEAVING

; TITLE OF INVENTION: SUBSTRATES CONTAINING DIBASIC RESIDUES

; NUMBER OF SEQUENCES: 74

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 1 DNA Way

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WinPatIn (Genentech)

; CURRENT APPLICATION DATA: US/08/398,028B

; FILING DATE: 03-Mar-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Kubinec, Jeffrey S.

; REGISTRATION NUMBER: 36,575

; REFERENCE/DOCKET NUMBER: P0936

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650/225-8228

; TELEFAX: 650/952-9881

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 8119 base pairs

; TYPE: Nucleic Acid

; STRANDEDNESS: Single

; TOPOLOGY: Linear

US-08-398-028B-1

Query Match 27.5%; Score 867; DB 2; Length 8119;

Best Local Similarity 100.0%; Pred. No. 2.8e-237; Indels 0; Gaps 0;

Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1475 TGCTTAGGAAGACGAGTTATTAATAGCTGAATAGAACGGTGTCTCTCAAAATATTTCTTAT 1534

Db 5764 TGCTTAGGAAGACGAGTTATTAATAGCTGAATAGAACGGTGTCTCTCAAAATATTTCTTAT 5823

Qy 1535 TTAGAAAAGCAATCTAAATATCTCAAAAGGAATGAGAAATAGTGAATGGACAATAA 1594

Db 5824 TTAGAAAAGCAATCTAAATATCTCAAAAGGAATGAGAAATAGTGAATGGACAATAA 5883

Qy 1595 TAATGACTAGAGAGAAAGAAATGAAGATTCCTCATGAAATTAAGGAACGATATTTGGATA 1654

Db 5884 TAATGACTAGAGAGAAAGAAATGAAGATTCCTCATGAAATTAAGGAACGATATTTGGATA 5943

Qy 1655 AATATGGGATGATGTTAAGGCTATTGGTGTATTTATGGCTCTCTTGGTCTGTCAGACTGATG 1714

Db 5944 AATATGGGATGATGTTAAGGCTATTGGTGTATTTATGGCTCTCTTGGTCTGTCAGACTGATG 6003

Qy 1715 GGCCCTATTCCGATATTGAGATGATGTCATGTCACAGAGAAAGCAGAGTTTCAGCC 1774

Db 6004 GGCCCTATTCCGATATTGAGATGATGTCATGTCACAGAGAAAGCAGAGTTTCAGCC 6063

Qy 1775 ATGATGGAACAACCGGTGAGTGGAAAGTGAAGTGAATTTTGTAGGAGAGAGATTTCTAC 1834

Db 6064 ATGATGGAACAACCGGTGAGTGGAAAGTGAAGTGAATTTTGTAGGAGAGAGATTTCTAC 6123

Qy 1835 TAGATTATGCTATCTCAGGTGGAATCAGATTGGCCGCTTTACACATGGTCAATTTTCTCTA 1894

Db 6124 TAGATTATGCTATCTCAGGTGGAATCAGATTGGCCGCTTTACACATGGTCAATTTTCTCTA 6183

Qy 1895 TTTTGGCGATTATGATTCAAGTGGATCTTATAGAGAAAGTGTATCAAACTCTCTAAATCGG 1954

Db 6184 TTTTGGCGATTATGATTCAAGTGGATCTTATAGAGAAAGTGTATCAAACTCTCTAAATCGG 6243

Qy 1955 TAGAAGCCCAACCGTTCCACGATGCGATTTCGTGCGCTTATCGTAGAGAGCTGTTTGAAT 2014

Db 6244 TAGAAGCCCAACCGTTCCACGATGCGATTTCGTGCGCTTATCGTAGAGAGCTGTTTGAAT 6303

Qy 2015 ATGCAGGCAAAATGCGGTAAATTCGTGTCGAAGACCGACAACTTTCTACCATCTTGA 2074

Db 6304 ATGCAGGCAAAATGCGGTAAATTCGTGTCGAAGACCGACAACTTTCTACCATCTTGA 6363

Qy 2075 CTGTACAGGTAGCAATGCGAGTGGCCATGTTGATTGGTCTGCATCATCGCATCTGTTATA 2134

Db 6364 CTGTACAGGTAGCAATGCGAGTGGCCATGTTGATTGGTCTGCATCATCGCATCTGTTATA 6423

Qy 2135 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTTATG 2194

Db 6424 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTTATG 6483

Qy 2195 ACCATCTGTGCCAGTTCGTAATGTCGTCAACTTTCCGACTCTGAGAACTTCTGGAAT 2254

Db 6484 ACCATCTGTGCCAGTTCGTAATGTCGTCAACTTTCCGACTCTGAGAACTTCTGGAAT 6543

Qy 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGACCGATATATAGTGG 2314

Db 6544 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGACCGATATATAGTGG 6603

Qy 2315 ATGTGTCAAAACGCATACCATTTTGAA 2341

Db 6604 ATGTGTCAAAACGCATACCATTTTGAA 6630

## RESULT 13

US-08-504-265B-1

; Sequence 1, Application US/08504265B

; Patent No. 5837516

; GENERAL INFORMATION:

; APPLICANT: Marcus D. Ballinger and James A. Wells

; TITLE OF INVENTION: SUBTILISIN VARIANTS CAPABLE OF CLEAVING  
; TITLE OF INVENTION: SUBSTRATES CONTAINING BASIC RESIDUES  
; NUMBER OF SEQUENCES: 90  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/504,265B  
; FILING DATE: 19-Jul-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/398028  
; FILING DATE: 03-MAR-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kubinec, Jeffrey S.  
; REGISTRATION NUMBER: 36,575  
; REFERENCE/DOCKET NUMBER: P0936P1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-8228  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 819 base pairs  
; TYPE: Nucleic Acid  
; STRANDEDNESS: Single  
; TOPOLOGY: Linear  
; US-08-504-265B-1

Query Match 27.5%; Score 867; DB 2; Length 819;  
Best Local Similarity 100.0%; Pred. No. 2.8e-237;  
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1475 TGGTTAGGAAGACAGTATTATAGCTGAATGAAGACGGTCTCTCCAAATATTTCTAT 1534  
DB 5764 TGGTTAGGAAGACAGTATTATAGCTGAATGAAGACGGTCTCTCCAAATATTTCTAT 5823  
QY 1535 TTAGAAAGCAATCTAAATTTATCTGAAAGGGAATGAATAGTGAATGGACCAATAA 1594  
DB 5824 TTAGAAAGCAATCTAAATTTATCTGAAAGGGAATGAATAGTGAATGGACCAATAA 5883  
QY 1595 TAATGACTAGAGAAGAAAGAAATGAAGATTCTTCATGAAATTAAGGAACGAATATTGGATA 1654  
DB 5884 TAATGACTAGAGAAGAAAGAAATGAAGATTCTTCATGAAATTAAGGAACGAATATTGGATA 5943  
QY 1655 AATATGGGATGATGTTAAGGCTATTGGTGTATTTATGGCTCTCTTGGTCTGACAGTATG 1714  
DB 5944 AATATGGGATGATGTTAAGGCTATTGGTGTATTTATGGCTCTCTTGGTCTGACAGTATG 6003  
QY 1715 GGCCCTATTCCGATATTGAGATGATGTCATGTCATGTCACAGAGGAGAGAGTTCAGCC 1774  
DB 6004 GGCCCTATTCCGATATTGAGATGATGTCATGTCATGTCACAGAGGAGAGAGTTCAGCC 6063  
QY 1775 ATGAATGGCAACACCGGTGAGTGAAGGTGAAGTGAATTTGATAGCGAAGAGATTTCTAC 1834  
DB 6064 ATGAATGGCAACACCGGTGAGTGAAGGTGAAGTGAATTTGATAGCGAAGAGATTTCTAC 6123  
QY 1835 TAGATTATGATCTCAGGTGGAATCAGATTGGCGCTTACATGATGTCATATTTTCTCTA 1894  
DB 6124 TAGATTATGATCTCAGGTGGAATCAGATTGGCGCTTACATGATGTCATATTTTCTCTA 6183  
QY 1895 TTTTGGCGATTTATGATTCAGGTGCGATCTAGGAAAGGTATCAAACTGCTAAATCGG 1954  
DB 6184 TTTTGGCGATTTATGATTCAGGTGCGATCTAGGAAAGGTATCAAACTGCTAAATCGG 6243

QY 1955 TAGAAGCCAAACGTTCCAGATGCGATTTGTGCCCTTATCTGTAAGAGAGCTGTTTGAAT 2014  
DB 6244 TAGAAGCCAAACGTTCCAGATGCGATTTGTGCCCTTATCTGTAAGAGAGCTGTTTGAAT 6303  
QY 2015 ATGACAGGCAAAATGCGTGAATATATTCGTGTCGAAGGACGCAACATTTCTACCATCCTTGA 2074  
DB 6304 ATGACAGGCAAAATGCGTGAATATATTCGTGTCGAAGGACGCAACATTTCTACCATCCTTGA 6363  
QY 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGGTCTGTCATCATCGCATCTGTTATA 2134  
DB 6364 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGGTCTGTCATCATCGCATCTGTTATA 6423  
QY 2135 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTATG 2194  
DB 6424 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTATG 6483  
QY 2195 ACCATCTGTCAGTTCGTAATGTCGTGTCACACTTTCCGACTCTGAGAACTTCTGGAAT 2254  
DB 6484 ACCATCTGTCAGTTCGTAATGTCGTGTCACACTTTCCGACTCTGAGAACTTCTGGAAT 6543  
QY 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACGATATATAGTGG 2314  
DB 6544 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACGATATATAGTGG 6603  
QY 2315 ATGTGTCAAAACGCATACCATTTTGA 2341  
DB 6604 ATGTGTCAAAACGCATACCATTTTGA 6630

## RESULT 14

US-08-556-978B-79/c  
; Sequence 79, Application US/08556978B  
; Patent No. 6268169  
; GENERAL INFORMATION:  
; APPLICANT: FAHNESTOCK, STEPHEN F.  
; TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED  
; NUMBER OF SEQUENCES: 107  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY  
; STREET: 1007 MARKET STREET  
; CITY: WILMINGTON  
; STATE: DELAWARE  
; COUNTRY: UNITED STATES OF AMERICA  
; ZIP: 19898  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: DISKETTE, 3.50 INCH  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: MICROSOFT WINDOWS 95  
; SOFTWARE: MICROSOFT WORD FOR WINDOWS 95  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/556,978B  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/077,600  
; FILING DATE: JUNE 15, 1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: FLOYD, LINDA AXAMETHY  
; REGISTRATION NUMBER: 33,692  
; REFERENCE/DOCKET NUMBER: CR-9389-A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 302-892-8112  
; TELEFAX: 302-773-0164  
; INFORMATION FOR SEQ ID NO: 79:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 9144 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: circular  
; MOLECULE TYPE: DNA (genomic)  
; US-08-556-978B-79

```
Query Match      27.5%; Score 867; DB 3; Length 9144;
Best Local Similarity 100.0%; Pred. No. 3e-237;
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1475 TGCTTAGGAACGAGGTTATTAAATAGCTGAATAGAACGGTGCTCTCCAAATATTCCTAT 1534
Db 7662 TGCTTAGGAACGAGGTTATTAAATAGCTGAATAGAACGGTGCTCTCCAAATATTCCTAT 7603

Qy 1535 TTAGAAAAGCAAACTCTAAAATTTATCTGAAAAGGGAATGAGAAATAGTGAATGGACCAATAA 1594
Db 7602 TTAGAAAAGCAAACTCTAAAATTTATCTGAAAAGGGAATGAGAAATAGTGAATGGACCAATAA 7543

Qy 1595 TAAATGACTAGAGAAGAAAGAAATGAAGATTGTTCATGAAATTTAAGGAACGAATATTTGGATA 1654
Db 7542 TAAATGACTAGAGAAGAAAGAAATGAAGATTGTTCATGAAATTTAAGGAACGAATATTTGGATA 7483

Qy 1655 AATATGGGGATGATGTTAAGGCTATTGGTGTATTGGCTCTCTTTGGTCTCAGACTGATG 1714
Db 7482 AATATGGGGATGATGTTAAGGCTATTGGTGTATTGGTCTCTCTTTGGTCTCAGACTGATG 7423

Qy 1715 GGCCCTATTTCGGATATTCAGATGATGCTGTCATGTCACAGAGGAAGCAGAGTTTCAGCC 1774
Db 7422 GGCCCTATTTCGGATATTCAGATGATGCTGTCATGTCACAGAGGAAGCAGAGTTTCAGCC 7363

Qy 1775 ATGAATGGACAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTAGTGAAGAGAGATTCTAC 1834
Db 7362 ATGAATGGACAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTAGTGAAGAGAGATTCTAC 7303

Qy 1835 TAGATTATGCAATCTCAGGTGAATCAGATTGGCGCTTTACACATGGTCAATTTTCTCTA 1894
Db 7302 TAGATTATGCAATCTCAGGTGAATCAGATTGGCGCTTTACACATGGTCAATTTTCTCTA 7243

Qy 1895 TTTTGGCCGATTTATGATTCAGTGGATCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 1954
Db 7242 TTTTGGCCGATTTATGATTCAGTGGATCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 7183

Qy 1955 TAGAAGCCCAAAACGGTTCACGATGCGATTTGTGCCCTTTATCGTAGAAGAGCTGTTTGAAT 2014
Db 7182 TAGAAGCCCAAAACGGTTCACGATGCGATTTGTGCCCTTTATCGTAGAAGAGCTGTTTGAAT 7123

Qy 2015 ATGCAGGCAAAATGGGTTAATTTGTGTGCAAGGACGGAACCAATTTCTACATCCTTGA 2074
Db 7122 ATGCAGGCAAAATGGGTTAATTTGTGTGCAAGGACGGAACCAATTTCTACATCCTTGA 7063

Qy 2075 CTGTACAGGTAGCAATGCGAGTGCCATGTTGATTTGTTGCTGCATCATCGCATCTGTTATA 2134
Db 7062 CTGTACAGGTAGCAATGCGAGTGCCATGTTGATTTGTTGCTGCATCATCGCATCTGTTATA 2003

Qy 2135 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCCTTCAGGTTATG 2194
Db 7002 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCCTTCAGGTTATG 6943

Qy 2195 ACCATCTGTGCGAGTTCTGTAATGTCGTCAACTTTCCGACTCTGAGAAACTTCTGGAAAT 2254
Db 6942 ACCATCTGTGCGAGTTCTGTAATGTCGTCAACTTTCCGACTCTGAGAAACTTCTGGAAAT 6883

Qy 2255 CGCTAGAGAAATTTCTGGAATGGATTTCAGGAGTGGACAGAACGACACGGATATATAGTGG 2314
Db 6882 CGCTAGAGAAATTTCTGGAATGGATTTCAGGAGTGGACAGAACGACACGGATATATAGTGG 6823

Qy 2315 ATGTGTCAAAACGCATACCATTTTGA 2341
Db 6822 ATGTGTCAAAACGCATACCATTTTGA 6796
```

RESULT 15  
5304637-21  
;Patent No. 5304637  
;APPLICANT: DORSSERS, LAMBERTUS C. J.; WAGMAKER, GERARD; VOS,  
;VONNE J.; VAN LEEN, ROBERT W.; PERSON; MARIA L. N.  
; TITLE OF INVENTION: EXPRESSION AND PURIFICATION OF HUMAN  
;INTERLEUKIN-3 AND MUTAINS THEREOF  
; NUMBER OF SEQUENCES: 30

```
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/494,182  
; FILING DATE: 13-MAR-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 249,184  
; FILING DATE: 16-AUG-1988  
; SEQ ID NO:21:  
; LENGTH: 7326  
5304637-21

Query Match      27.4%; Score 865.4; DB 9; Length 7326;  
Best Local Similarity 99.9%; Pred. No. 7.6e-237;  
Matches 866; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1475 TGCTTAGGAACGAGGTTATTAAATAGCTGAATAGAACGGTGCTCTCCAAATATTCCTAT 1534
Db 5331 TGCTTAGGAACGAGGTTATTAAATAGCTGAATAGAACGGTGCTCTCCAAATATTCCTAT 5290

Qy 1535 TTAGAAAAGCAAACTCTAAAATTTATCTGAAAAGGGAATGAGAAATAGTGAATGGACCAATAA 1594
Db 5291 TTAGAAAAGCAAACTCTAAAATTTATCTGAAAAGGGAATGAGAAATAGTGAATGGACCAATAA 5350

Qy 1595 TAAATGACTAGAGAAGAAAGAAATGAAGATTGTTCATGAAATTTAAGGAACGAATATTTGGATA 1654
Db 5351 TAAATGACTAGAGAAGAAAGAAATGAAGATTGTTCATGAAATTTAAGGAACGAATATTTGGATA 5410

Qy 1655 AATATGGGATGATGTTAAGGCTATTGGTGTATTGGCTCTCTTTGGTCTCAGACTGATG 1714
Db 5411 AATATGGGATGATGTTAAGGCTATTGGTGTATTGGCTCTCTTTGGTCTCAGACTGATG 5470

Qy 1715 GGCCCTATTTCGGATATTCAGATGATGCTCAACAGAGGAAGCAGAGTTTCAGCC 1774
Db 5471 GGCCCTATTTCGGATATTCAGATGATGCTCAACAGAGGAAGCAGAGTTTCAGCC 5530

Qy 1775 ATGAATGGACAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTAGTGAAGAGAGATTCTAC 1834
Db 5531 ATGAATGGACAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTAGTGAAGAGAGATTCTAC 5590

Qy 1835 TAGATTATGCAATCTCAGGTGAATCAGATTGGCGCTTTACACATGGTCAATTTTCTCTA 1894
Db 5591 TAGATTATGCAATCTCAGGTGAATCAGATTGGCGCTTTACACATGGTCAATTTTCTCTA 5650

Qy 1895 TTTTGGCCGATTTATGATTCAGGTGGATACTTAGAGAAAGTGTATCAAACTGCTAAATCGG 1954
Db 5651 TTTTGGCCGATTTATGATTCAGGTGGATACTTAGAGAAAGTGTATCAAACTGCTAAATCGG 5710

Qy 1955 TAGAAGCCCAAAACGGTTCACGATGCGATTTGTGCCCTTTATCGTAGAAGAGCTGTTTGAAT 2014
Db 5711 TAGAAGCCCAAAACGGTTCACGATGCGATTTGTGCCCTTTATCGTAGAAGAGCTGTTTGAAT 5770

Qy 2015 ATGCAGGCAAAATGGCGTAAATTTCTGTCGAAAGGACGGAACCAATTTCTACATCCTTGA 2074
Db 5771 ATGCAGGCAAAATGGCGTAAATTTCTGTCGAAAGGACGGAACCAATTTCTACATCCTTGA 5830

Qy 2075 CTGTACAGGTAGCAATGCGAGTGCCATGTTGATTTGTTGCTGCATCATCGCATCTGTTATA 2134
Db 5831 CTGTACAGGTAGCAATGCGAGTGCCATGTTGATTTGTTGCTGCATCATCGCATCTGTTATA 5890

Qy 2135 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCCTTCAGGTTATG 2194
Db 5891 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCCTTCAGGTTATG 5950

Qy 2195 ACCATCTGTGCGAGTTCTGTAATGTCGTCAACTTTCCGACTCTGAGAAACTTCTGGAAAT 2254
Db 5951 ACCATCTGTGCGAGTTCTGTAATGTCGTCAACTTTCCGACTCTGAGAAACTTCTGGAAAT 6010

Qy 2255 CGCTAGAGAAATTTCTGGAATGGGATTTCAGGAGTGGACAGAACGACACGGATATATAGTGG 2314
Db 6011 CGCTAGAGAAATTTCTGGAATGGGATTTCAGGAGTGGACAGAACGACACGGATATATAGTGG 6070

Qy 2315 ATGTGTCAAAACGCATACCATTTTGA 2341
Db 6071 ATGTGTCAAAACGCATACCATTTTGA 6097
```

Search completed: February 11, 2006, 23:26:19  
Job time : 403 secs

***This Page Blank (uspto)***



Qy	301	GCCAGCTTTAAACTGACAGAAAGCGCCCTGCTGTTTTTCGAGCGGTACTTTCGSCCAATGTC	360
Dd	301		
Qy	361	GGTGCTTTTCAATCCTTCCAGAAAGGAAGATGTCATTTTAAAGTGACCACTCAATCAT	420
Dd	361		
Qy	421	GCAAGTATGATCGAGCGCTGCGGACTTTCTTAAGGCTGATACAGTTGTTTATTCGGCATATT	480
Dd	421		
Qy	481	GATATGAATGATCTTGAATAACAGCTGAATGAACAACACAGCGTTATCAGCGCCGTTTTATC	540
Dd	481		
Qy	541	GTAACAGACGGAGTATTCAGCATGGAATGGAACAATCGCCCTCTTTGATCAGATCATCTCA	600
Dd	541		
Qy	601	CTTGGAAACGCTATCATGCTTGGTGTGTTGATGATGCCACGCAACAGGAGTTTTTG	660
Dd	601		
Qy	661	GGCGATTCCGGACAAAGGAACGAGTGAATACATTTTGGTGTGTTGTCGCCGACATTTGTTATCGGC	720
Dd	661		
Qy	721	ACCTTAAAGCAAGCTGTTGGCGGGAAGGAGTTTTTGGCGAGGATCAGCGGTCTTCATC	780
Dd	721		
Qy	781	GACTTTTGTGTAACCATGCCAGAACATTTATCTTTTAAACGGCTATTCGCCACGCGACG	840
Dd	781		
Qy	841	TGTCGGCTGCTCAGAGGCTTTTCAACATCATTTGAAGCCAGCAGGAAACAGCAGCTT	900
Dd	841		
Qy	901	TTATTTTCTTATATCAGCATGATCAGAACCAAGTCTGAAGAAATATGGGTATATGGTGA	960
Dd	901		
Qy	961	GGAGATCACACCGATTTATTCCTGTAGTCAITGGCGATGCCATATAAACCGTCTATTT	1020
Dd	961		
Qy	1021	GCTGAAACCTGCAAGGCAAGGAATTTATGCTCTGCTCATTTGCGCGCCCAACCGTTGCG	1080
Dd	1021		
Qy	1081	CCGGGTGAAAGCCGGATTCGAAGCTTGGCGACGAGGTGCGATCAGGGAATGAGTTTATA	1140
Dd	1081		
Qy	1141	AAATATAAAGCACTGAAAAAGGTGCTTTTTTTTGTAGTGGTTTGAACCTGTTCTTTCT	1200
Dd	1141		
Qy	1201	ATCTTGATACATATAGAAATAACGTCAATTTTATTTTATTTTATTTAGTTGCTGAAAGGTGCG	1260
Dd	1201		
Qy	1261	TTGAAGTGTGGTATGATGTGTTTTTAAAGTATTGAAACCCCTTAAATTTGGTTGCAAG	1320
Dd	1261		
Qy	1321	AAAAACCCCATCTGTTAAAGTTTATAAGTGACTAAACAAATAAATAAGATAGATGGGGTTT	1380
Dd	1321		

Qy	1381	CTTTTAATATTATGTGCTCTAATAGTAGCATTTTATTTCAGATGAAAAATCAAGGGTTTTAG	1440
Dd	1381		
Qy	1441	TGGACAAAGCAAAAGTGGAAAAAGTGAGACCATGTGCTTAGGAAAGACGAGTTATTAATAG	1500
Dd	1441		
Qy	1501	CTGNAATGAAGAACGGTGTCTCCAAATATTCTTATTTTAGAAGCAAAATCTAAAAATATCT	1560
Dd	1501		
Qy	1561	GAAAAAGGAATGAGAAATAGTGAATGGACCAATATAATAGTACAGAGAAAGAAATGAAG	1620
Dd	1561		
Qy	1621	ATTGTTTCATGAAATTAAGGAACGAATATTGGAATAAATATGGGGATGATGTTTAAAGGCTATT	1680
Dd	1621		
Qy	1681	GGTGTGTTTATGGCTCTCTTGGTCTGACACTGATGGGCCCTATTTCGGATATTGAGATGATG	1740
Dd	1681		
Qy	1741	TGTGTCATGTCAACAGAGGAAGCAGAGTTTCAAGCATGAATGGACAAACCGGTGAGTGGAAAG	1800
Dd	1741		
Qy	1801	GTGGAAGTGAATTTTGTATAGCAAGAGATTTCTACTAGATTATGCAATCTCAGGTGGAATCA	1860
Dd	1801		
Qy	1861	GATTGGCCGCTTACACATGCTCAATTTTCTCTATTTTTCGGATTTATGATTCAGGTGGA	1920
Dd	1861		
Qy	1921	TACTTAGAAGAAAGTGTATCAAACTGCTAAATCGGTAGAACCCCAAGCTTCCACGATGCG	1980
Dd	1921		
Qy	1981	ATTGTGCCCTTATCGTAGAAGAGCTGTTTGAATATATCGAGCAAAATGGCGTAAATTTCTGT	2040
Dd	1981		
Qy	2041	GTGCAAGGACCGCAACATTTTACCATCTGCTGCTACAGGTAGCAATGGCAGGTGCC	2100
Dd	2041		
Qy	2101	ATGTTGATTGGTCTGCAATCATCGCATCTGTTATACGACGAGCGCTTCGGTCTTAACTGAA	2160
Dd	2101		
Qy	2161	GCAGTTAAGCAATCAGATCTTCTTTCAGGTTATGACCAATCTGTGCCAGTTCTGTAATGTCT	2220
Dd	2161		
Qy	2221	GGTCAACTTTCCGACTCTGAGAACTTCTCGAATCGCTAGAGATTTTCTGGAATGGGATTT	2280
Dd	2221		
Qy	2281	CAGGAGTGGACAGAAACGACACGGATATATAGTGGATGTGTCAAAACGCATACCACTTTTGA	2340
Dd	2281		
Qy	2341	ATTTCGAAGCGCGGATTTGAGTCTTTACCGGATGGTGAATAAGAAACGCTGCTTGAAGGCG	2400
Dd	2341		
Qy	2401	CGAAGCGGGCGCAGATCTGAATATCGGCACATATTGATCGTGGCAAGCGCGAGAGGTC	2460
Dd	2401		
Qy	2461	CGTCTAACAGAGAAGTGGATCAGGTCGTAGATCGGTTTCAGGNAATTTAAAGAGACGTATG	2520



Db	2461	CGTCTAACAGAGAAGTGGATCAGGTCTAGATCGGTTCAGGAAATTAAGAGACGCTATG	2520
Qy	2521	GACTGAGATTTGTCATGCTCTGGACTGTTGAAGCCAGAGCAGGCGAAGCGCTCAAAG	2580
Db	2521	GACTGAGATTTGTCATGCTCTGGACTGTTGAAGCCAGAGCAGGCGAAGCGCTCAAAG	2580
Qy	2581	ATGCAGAGTAGACCGCTATAATCATATTTGAATACGTACAGAGAAACCATTCAAACA	2640
Db	2581	ATGCAGAGTAGACCGCTATAATCATATTTGAATACGTACAGAGAAACCATTCAAACA	2640
Qy	2641	TCACAACCTCACATACATACATACATACATACATACATACATACATACATACATACAT	2700
Db	2641	TCACAACCTCACATACATACATACATACATACATACATACATACATACATACATACAT	2700
Qy	2701	GGCTGCTCGGTGTTCAAGCGCCATATTCGGATGGAAGGAGCGAACAGGATGTCATTG	2760
Db	2701	GGCTGCTCGGTGTTCAAGCGCCATATTCGGATGGAAGGAGCGAACAGGATGTCATTG	2760
Qy	2761	ACATCGCAAAAGCTTGAAGGCTCTTGACCGGATTCATTCCTGTGAAATTTTTCATG	2820
Db	2761	ACATCGCAAAAGCTTGAAGGCTCTTGACCGGATTCATTCCTGTGAAATTTTTCATG	2820
Qy	2821	CAATTGATGGAGCGCCTTGAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGG	2880
Db	2821	CAATTGATGGAGCGCCTTGAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGG	2880
Qy	2881	TGCTGGCGCTGTTCCGCTTATCAATCCATCAAGGAAATTCGATTTCCGGAGGAGAG	2940
Db	2881	TGCTGGCGCTGTTCCGCTTATCAATCCATCAAGGAAATTCGATTTCCGGAGGAGAG	2940
Qy	2941	AGGTCAATCTCCGACATTCGAGGCTTACGGCTTTACGGCGCAAACTCCATTTTGTGCG	3000
Db	2941	AGGTCAATCTCCGACATTCGAGGCTTACGGCTTTACGGCGCAAACTCCATTTTGTGCG	3000
Qy	3001	GAGACTACTTAAACAACTGCGGCGCAAGAGGAGCGGAGGATCATATAATGCTGAGTGATT	3060
Db	3001	GAGACTACTTAAACAACTGCGGCGCAAGAGGAGCGGAGGATCATATAATGCTGAGTGATT	3060
Qy	3061	TAGCTTTGAAGTTGAATCAGTGAAGAAATGAAGGCTAGTTAAAGTGCAGAAAGCTGAA	3120
Db	3061	TAGCTTTGAAGTTGAATCAGTGAAGAAATGAAGGCTAGTTAAAGTGCAGAAAGCTGAA	3120
Qy	3121	AGATCAATAAAGCAATCGGTATGATGTCGAATTC	3156
Db	3121	AGATCAATAAAGCAATCGGTATGATGTCGAATTC	3156

RESULT 2

US-10-505-855-1  
; Sequence 1, Application US/10505855  
; Publication No. US20050124030A1  
; GENERAL INFORMATION:  
; APPLICANT: Roche Vitamins AG  
; TITLE OF INVENTION: Fermentation Process  
; FILE REFERENCE: 21154  
; CURRENT APPLICATION NUMBER: US/10/505,855  
; CURRENT FILING DATE: 2004-08-26  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 3560  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)..(3560)  
; OTHER INFORMATION: plasmid pNMR21  
US-10-505-855-1

Query Match 38.6%; Score 1219; DB 9; Length 3560;  
Best Local Similarity 99.5%; Pred. No. 1.3e-295;  
Matches 1235; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

Qy	1102	AGCTTGGGCGAGCGTTCGAGATCAGGGAATGAGTTTATAAAATAAAAAACGACCTGAAA	1161
Db	1466	AGCTTGGGCGAGCGTTCGAGATCAGGGAATGAGTTTATAAAATAAAAAACGACCTGAAA	1525
Qy	1162	AGGTGCTCTTTTGTGATGGTTTGAACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1221
Db	1526	AGGTGCTCTTTTGTGATGGTTTGAACCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1585
Qy	1222	ACGTCAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT	1281
Db	1586	ACGTC-----ATTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT	1639
Qy	1282	GTTTTAAAGTATTGAAACCCCTTTAAATTTGGTTTGCAGAAAAACCCCATCTGTAAAGT	1341
Db	1640	GTTTTAAAGTATTGAAACCCCTTTAAATTTGGTTTGCAGAAAAACCCCATCTGTAAAGT	1699
Qy	1342	TATAAGTACTAAACAAATAAATAAATAGATGGGGTTCCTTTTAAATATTATTTATGTCCTA	1401
Db	1700	TATAAGTACTAAACAAATAAATAAATAGATGGGGTTCCTTTTAAATATTATTTATGTCCTA	1759
Qy	1402	ATAGTAGCATTTTATTCAGATGAAAAATCAAGGGTTTTAGTGGCAAGCAAAAAAGTCGAA	1461
Db	1760	ATAGTAGCATTTTATTCAGATGAAAAATCAAGGGTTTTAGTGGCAAGCAAAAAAGTCGAA	1819
Qy	1462	AACTGAGACCATGCTGCTTAGGAAAGACGAGTTATTAATAGCTGAATAAGAACGGTCTCTC	1521
Db	1820	AACTGAGACCATGCTGCTTAGGAAAGACGAGTTATTAATAGCTGAATAAGAACGGTCTCTC	1879
Qy	1522	CAATATTTCTTATTTAGAAAGCAAAATCTAAATTTATCTGAAAGGGAATGAGAAATAGTG	1581
Db	1880	CAATATTTCTTATTTAGAAAGCAAAATCTAAATTTATCTGAAAGGGAATGAGAAATAGTG	1939
Qy	1582	AATGGACCAATAATATGACTAGAGAGAAAGAAATGAAGATTTCTCATGAAATTAAGGAA	1641
Db	1940	AATGGACCAATAATATGACTAGAGAGAAAGAAATGAAGATTTCTCATGAAATTAAGGAA	1999
Qy	1642	CGAATATTGGATAAATATGGGATGATTTAAGGCTATTGGTGTATTATGGCTCTCTTTGGT	1701
Db	2000	CGAATATTGGATAAATATGGGATGATTTAAGGCTATTGGTGTATTATGGCTCTCTTTGGT	2059
Qy	1702	CGTCAGATGATGGGCGCTTATTCGGATATTGAGATGATGATGATGATGATGATGATGATG	1761
Db	2060	CGTCAGATGATGGGCGCTTATTCGGATATTGAGATGATGATGATGATGATGATGATGATG	2119
Qy	1762	GCAGAGTTTCAAGCCATGAATGGAACAACCGGTGAGTGAAGGAGTGAAGTGAAGTGAAGT	1821
Db	2120	GCAGAGTTTCAAGCCATGAATGGAACAACCGGTGAGTGAAGGAGTGAAGTGAAGTGAAGT	2179
Qy	1822	GAAGAGATTTCTACTAGATTTATGATCTCAGGTGGAATCAGATTGGCGCTTTACACATGGT	1881
Db	2180	GAAGAGATTTCTACTAGATTTATGATCTCAGGTGGAATCAGATTGGCGCTTTACACATGGT	2239
Qy	1882	CAATTTTCTCTATTTTGGCGATTTATGATTCAGGTGGAATCATTAGAGAAAGTGTATCAA	1941
Db	2240	CAATTTTCTCTATTTTGGCGATTTATGATTCAGGTGGAATCATTAGAGAAAGTGTATCAA	2299
Qy	1942	ACTGCTAAATCGGTAGAACCCCAACGTTCCAGATGCGATTTGTGCCCTTATCGTAGAA	2001
Db	2300	ACTGCTAAATCGGTAGAACCCCAACGTTCCAGATGCGATTTGTGCCCTTATCGTAGAA	2359
Qy	2002	GAGCTGTTTGAATATGAGGCAAAATGGCGTAATATTCTGCTGCAAGGACCGCAACAATTT	2061
Db	2360	GAGCTGTTTGAATATGAGGCAAAATGGCGTAATATTCTGCTGCAAGGACCGCAACAATTT	2419
Qy	2062	CTACCATCTTGTGATGATGAGGAGCAATGCGAGGTGCGCATGTTGATTTGTTGTCATCAT	2121
Db	2420	CTACCATCTTGTGATGATGAGGAGCAATGCGAGGTGCGCATGTTGATTTGTTGTCATCAT	2479
Qy	2122	CGCATCTGTTATACGAGCGCTTCGGTCTTAACTGAACGAGTGAAGCAATCAGATCTT	2181
Db	2480	CGCATCTGTTATACGAGCGCTTCGGTCTTAACTGAACGAGTGAAGCAATCAGATCTT	2539

Qy 2182 CTTTCAGGTTATGACCATCTGTGCAGTTTCGTAATGTCGTCAACTTTTCGACTCTGAG 2241  
Db 2540 CTTTCAGGTTATGACCATCTGTGCAGTTTCGTAATGTCGTCAACTTTTCGACTCTGAG 2599  
Qy 2242 AAATCTTCGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGACAGAACAC 2301  
Db 2600 AAATCTTCGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGACAGAACAC 2659  
Qy 2302 GGATATATAGTGGATGTGTCAAAAGCATACATTTTGAAT 2342  
Db 2660 GGATATATAGTGGATGTGTCAAAAGCATACATTTTGAAT 2700

RESULT 3  
US-10-369-493-46942  
; Sequence 46942, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 46942  
; LENGTH: 1170  
; TYPE: DNA  
; ORGANISM: Bacillus subtilis  
US-10-369-493-46942

Query Match 34.3%; Score 1081; DB 6; Length 1170;  
Best Local Similarity 100.0%; Pred. No. 3.9e-461;  
Matches 1081; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 22 TTGAAGATTGATTCCTCGTTAAACGAGCGGTTAGACAGAAATGAAAGAGCGCGGTACAT 81  
Db 1 TTGAAGATTGATTCCTCGTTAAACGAGCGGTTAGACAGAAATGAAAGAGCGCGGTACAT 60

Qy 82 CGTAACCTGCGGTCAATGGATGAGAGCGCGGTTCCAGAGAGAAATATGATGGCGAAAT 141  
Db 61 CGTAACCTGCGGTCAATGGATGAGAGCGCGGTTCCAGAGAGAAATATGATGGCGAAAT 120

Qy 142 CAACGGTCTCGTCTCAACAATTTATTTAGGGCTCGCAAGCGATAGACGTTTCATCGAT 201  
Db 121 CAACGGTCTCGTCTCAACAATTTATTTAGGGCTCGCAAGCGATAGACGTTTCATCGAT 180

Qy 202 GCAGCCCAACAGCATTCGACAAATTTGGGACAGAGACGCGGTTACGTTTAAACGACA 261  
Db 181 GCAGCCCAACAGCATTCGACAAATTTGGGACAGAGACGCGGTTACGTTTAAACGACA 240

Qy 262 GGCAATTCGGTCTGGCATGAAAGCTAGAAAGAGATTCGACGTTTAAACTGACAGAA 321  
Db 241 GGCAATTCGGTCTGGCATGAAAGCTAGAAAGAGATTCGACGTTTAAACTGACAGAA 300

Qy 322 GCGGCCCTGCTGTTTTCGAGCGGTACTTTGGCCAAATGTCGGTGTCCTTCATCTGCCA 381  
Db 301 GCGGCCCTGCTGTTTTCGAGCGGTACTTTGGCCAAATGTCGGTGTCCTTCATCTGCCA 360

Qy 382 GAAAGGAAGATGTCAATTTTAAGTGACCGAGTCAATCATGCAAGTATGATCGACGGTGC 441  
Db 361 GAAAGGAAGATGTCAATTTTAAGTGACCGAGTCAATCATGCAAGTATGATCGACGGTGC 420

Qy 442 CGACTTTCTAAGCGTGATACAGTTGTTTATCGGCATATTTGATGATGATGATGATGATGAT 501  
Db 421 CGACTTTCTAAGCGTGATACAGTTGTTTATCGGCATATTTGATGATGATGATGATGATGAT 480

Qy 502 AAGCTGAATGAAACACACAGCGGTTATCAGCGCGGTTTTATCGTAACACAGCGGATTTACAGC 561  
Db 481 AAGCTGAATGAAACACACAGCGGTTATCAGCGCGGTTTTATCGTAACACAGCGGATTTACAGC 540  
Qy 562 ATGGATGGCAAAATCGCCCTCTTTGATCAGATCATCTCACTTGGGAAACGCTATCATGCC 621  
Db 541 ATGGATGGCAAAATCGCCCTCTTTGATCAGATCATCTCACTTGGGAAACGCTATCATGCC 600  
Qy 622 TTCTGTGTCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 681  
Db 601 TTCTGTGTCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660  
Qy 682 AGTGAATACTTTTGGTGTTCGCGACATTTGTTATCGGCACCTTAAGCAAAAGCTGTTGGC 741  
Db 661 AGTGAATACTTTTGGTGTTCGCGACATTTGTTATCGGCACCTTAAGCAAAAGCTGTTGGC 720  
Qy 742 GCGGAAGGAGGTTTTCGCGAGGATCAGCGGTCTTCATCGACTTTTTCGCTGAAACCATGCC 801  
Db 721 GCGGAAGGAGGTTTTCGCGAGGATCAGCGGTCTTCATCGACTTTTTCGCTGAAACCATGCC 780  
Qy 802 AGAATATTATCTTTGAAACGCGCTATTCGCCACGAGCTGTGGGCTGCTCAGAGGCT 861  
Db 781 AGAATATTATCTTTGAAACGCGCTATTCGCCACGAGCTGTGGGCTGCTCAGAGGCT 840  
Qy 862 TTCAACATCATTTGAAGCAGCAGGAAACACAGAGCTTTTATTTCTTATATCAGCATG 921  
Db 841 TTCAACATCATTTGAAGCAGCAGGAAACACAGAGCTTTTATTTCTTATATCAGCATG 900  
Qy 922 ATCAGAACCATCTGGAAGAAATATGGGTTATGTGTGAAAGAGATCACACACCATTTAT 981  
Db 901 ATCAGAACCATCTGGAAGAAATATGGGTTATGTGTGAAAGAGATCACACACCATTTAT 960  
Qy 982 CCTGTAGTCAATTGCGGATGCCATAAAACCGTCTCTATTTCTGTAAGAACTGCAGGGCAAG 1041  
Db 961 CCTGTAGTCAATTGCGGATGCCATAAAACCGTCTCTATTTCTGTAAGAACTGCAGGGCAAG 1020  
Qy 1042 GGAATTTATGCTCTGCGCAATTCGCGCCGCAACCGGTTGCGCGGTTGAAAGCGGATTCGA 1101  
Db 1021 GGAATTTATGCTCTGCGCAATTCGCGCCGCAACCGGTTGCGCGGTTGAAAGCGGATTCGA 1080  
Qy 1102 A 1102  
Db 1081 A 1081

RESULT 4  
US-09-869-855A-3  
; Sequence 3, Application US/09869855A  
; Publication No. US20030044940A1  
; GENERAL INFORMATION:  
; APPLICANT: Rasmussen, Michael Dolberg  
; TITLE OF INVENTION: Method For Increasing Gene Copy Number  
; FILE REFERENCE: 10028.204-US  
; CURRENT APPLICATION NUMBER: US/09/869,855A  
; CURRENT FILING DATE: 2001-07-05  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 5793  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic  
US-09-869-855A-3

Query Match 27.5%; Score 867; DB 3; Length 5793;  
Best Local Similarity 100.0%; Pred. No. 7.4e-207;  
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1475 TGCTTAGGAAGACAGGTTTATTAATAGCTGAATGAACGGTGTCTCTCAAAATATCTTAT 1534  
Db 968 TGCTTAGGAAGACAGGTTTATTAATAGCTGAATGAACGGTGTCTCTCAAAATATCTTAT 1027

Qy	1535	TTAGAAAAGC	AAATCTAA	AAATTA	CTGAAA	AGGGA	TAGAA	TAGTGA	ATGG	ACCA	TAA	1594
Db	1028	TTGAAAAGC	AAATCTAA	AAATTA	CTGAAA	AGGGA	TGGA	ATAGTGA	ATGG	ACCA	TAA	1087
Qy	1595	TAATGACT	TAGAGA	AAAGAA	TGAAGAT	TGTT	CAT	TGAAA	TTAAGGA	ACGA	ATATTGG	1654
Db	1088	TNATGACT	TAGAGA	AAAGAA	TGAAGAT	TGTT	CAT	TGAAA	TTAAGGA	ACGA	ATATTGG	1147
Qy	1655	AAATATGG	GGATGAT	TTTAAG	CGCTAT	TTCGGT	TTAT	TGGCT	CTCTTTGG	TCGT	CAGAC	1714
Db	1148	AAATATGG	GGATGAT	TTTAAG	CGCTAT	TTCGGT	TTAT	TGGCT	CTCTTTGG	TCGT	CAGAC	1207
Qy	1715	GGCCCTAT	TTCGGAT	TATTCAG	ATGAT	TGTCAT	GTCAA	CAGAGGA	AGCAG	AGTTTCAG	CC	1774
Db	1208	GGCCCTAT	TTCGGAT	TATTCAG	ATGAT	TGTCAT	GTCAA	CAGAGGA	AGCAG	AGTTTCAG	CC	1267
Qy	1775	ATGAAT	CGACA	CCGGT	CGAGT	TGGAAG	TGAA	TTTTTG	TATG	PAGCGAA	GAGATCT	1834
Db	1268	ATGAAT	CGACA	CCGGT	CGAGT	TGGAAG	TGAA	TTTTTG	TATG	PAGCGAA	GAGATCT	1327
Qy	1835	TAGATTA	TGAT	CTCAG	GTGGA	TACAG	ATTTGG	CGCGCT	TACAT	TGTC	CAATTTTTCT	1894
Db	1328	TAGATTA	TGAT	CTCAG	GTGGA	TACAG	ATTTGG	CGCGCT	TACAT	TGTC	CAATTTTTCT	1387
Qy	1895	TTTTTGG	CCGATTTAT	TGATTT	CAGGT	GGAT	CTATTT	TAGAA	GAAGTG	TATCAA	ACTGCT	1954
Db	1388	TTTTTGG	CCGATTTAT	TGATTT	CAGGT	GGAT	CTATTT	TAGAA	GAAGTG	TATCAA	ACTGCT	1447
Qy	1955	TAGAAG	CCCCAAA	CGTTCC	ACGAT	GCAG	ATTTGT	GTGCC	TTATCG	TAGA	AGAGCT	2014
Db	1448	TAGAAG	CCCCAAA	CGTTCC	ACGAT	GCAG	ATTTGT	GTGCC	TTATCG	TAGA	AGAGCT	1507
Qy	2015	ATGAGG	CAAAATGG	CGTA	ATAATTC	GTGTG	CAAGG	ACGACA	CAATTTCT	TACCA	CTCTTG	2074
Db	1508	ATGAGG	CAAAATGG	CGTA	ATAATTC	GTGTG	CAAGG	ACGACA	CAATTTCT	TACCA	CTCTTG	1567
Qy	2075	CTGTAC	AGGTAG	CAATGG	CAAGT	GCCAT	TTTGA	TTGGT	CTG	GCATCAT	CGCAT	2134
Db	1568	CTGTAC	AGGTAG	CAATGG	CAAGT	GCCAT	TTTGA	TTGGT	CTG	GCATCAT	CGCAT	1627
Qy	2135	CGACG	AGCGTT	CGGTC	TTAACT	GAAG	CAAGT	TAA	GC	AAATCAG	ATCTTC	2194
Db	1628	CGACG	AGCGTT	CGGTC	TTAACT	GAAG	CAAGT	TAA	GC	AAATCAG	ATCTTC	1687
Qy	2195	ACCAT	CTG	CGCAG	TTGCT	TAATG	CT	CGGT	CNA	CTTTTCC	GACTCT	2254
Db	1688	ACCAT	CTG	CGCAG	TTGCT	TAATG	CT	CGGT	CNA	CTTTTCC	GACTCT	1747
Qy	2255	CGCT	TAGAGA	ATTTCT	GGAAT	TGGGAT	TTCAG	GAGT	TGGA	CAGAA	CGAC	2314
Db	1748	CGCT	TAGAGA	ATTTCT	GGAAT	TGGGAT	TTCAG	GAGT	TGGA	CAGAA	CGAC	1807
Qy	2315	ATGTGT	CAAAA	CGCAT	ACCA	TTTTTG	AA					2341
Db	1808	ATGTGT	CAAAA	CGCAT	ACCA	TTTTTG	AA					1834

## RESULT 5

US-10-402-466A-2  
; Sequence 2, Application US/10402466A  
; Publication No. US20040028695A1  
; GENERAL INFORMATION:  
; APPLICANT: Park, Sukjoon  
; APPLICANT: Giri, Lallan  
; TITLE OF INVENTION: RECOMBINANT IMMUNOGENIC COMPOSITIONS AND METHODS FOR PROTECTING  
; TITLE OF INVENTION: AGAINST LETHAL INFECTIONS FROM BACILLUS ANTHRACIS  
; FILE REFERENCE: 18933-00005  
; CURRENT APPLICATION NUMBER: US/10/402,466A  
; CURRENT FILING DATE: 2003-03-20  
; PRIOR APPLICATION NUMBER: 60/372,152  
; PRIOR FILING DATE: 2002-04-12  
; NUMBER OF SEQ ID NOS: 34  
; SEQ ID NO 2

	;	LENGTH: 5865	
	;	TYPE: DNA	
	;	ORGANISM: Artificial	
	;	FEATURE:	
	;	NAME/KEY:	
	;	LOCATION:	
	;	OTHER INFORMATION: Artificial DNA sequence to be used as one of the two backbones.	
	;	OTHER INFORMATION: E. coli-Bacillus anthracis shuttle vectors. Designated US-10-402-466A-2	
		Query Match	
		Best Local Similarity 100.0%;	Pred. No. 7.4e-207; Length 5865;
		Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1475	TGCTTAGGAGACGAGTTATTAATAGCTGAATAAGAACCGGTGCTCTCCAAATATTCTTAT	1534
Dd	4390	TGCTTAGGAGACGAGTTATTAATAGCTGAATAAGAACCGGTGCTCTCCAAATATTCTTAT	4449
Qy	1535	TTAGAAAAGCAATCTAAATTTATCTGAAAAGGGATAGAAATAGTGAAATGGACCAATAA	1594
Dd	4450	TTAGAAAAGCAAATCTAAATTTATCTGAAAAGGGAAATAGAAATAGTGAAATGGACCAATAA	4509
Qy	1595	TAATGACTAGAGAAGAAAGAACTCAAGATTGTTTCATGAAATTAAGGAACGAAATTTGGATA	1654
Dd	4510	TAATGACTAGAGAAGAAAGAAATGAAGATTGTTTATGAAATTAAGGAACGAAATTTGGATA	4569
Qy	1655	AATATGGGATGATGTTAAGGCCTATTGGTGTATTGGCTCTCTTTGGTCTGTCAGACTGATG	1714
Dd	4570	AATATGGGATGATGTTAAGGCCTATTGGTGTATTGGCTCTCTTTGGTCTGTCAGACTGATG	4629
Qy	1715	GGCCCTATTCCGANAATTAGATGATGTGTGTCAATGTCACAAGAGGAAGCAGAGTTTCAGCC	1774
Dd	4630	GGCCCTATTCCGANAATTAGATGATGTGTGTCAATGTCACAAGAGGAAGCAGAGTTTCAGCC	4689
Qy	1775	ATGAATGGACAAACCGGTGAGTGGAAAGTGGAAAGTGAATTTTGATAGCGAAGAGATTCTAC	1834
Dd	4690	ATGAATGGACAAACCGGTGAGTGGAAAGTGGAAAGTGAATTTTGATAGCGAAGAGATTCTAC	4749
Qy	1835	TAGATTATGCATCTCAGGTGGAATCAGATTGGCGCGCTTACACATGGTCAATTTTTCTCTA	1894
Dd	4750	TAGATTATGCATCTCAGGTGGAATCAGATTGGCGCGCTTACACATGGTCAATTTTTCTCTA	4809
Qy	1895	TTTTGCCGATTTATGATTTCAAGTGGATCATTTAGAGAAAGTGTATCAAACTGCTAAATFCGG	1954
Dd	4810	TTTTGCCGATTTATGATTTCAAGTGGATCATTTAGAGAAAGTGTATCAAACTGCTAAATFCGG	4869
Qy	1955	TAGAGCCCAAAACGTTCCACGATGCGATTTGTGTCCCTTATCGTAGAAGAGCTGTTGGAAT	2014
Dd	4870	TAGAGCCCAAAACGTTCCACGATGCGATTTGTGTCCCTTATCGTAGAAGAGCTGTTGGAAT	4929
Qy	2015	ATCAGGCAAAATGGCGTAATATTTCGTGTGCAAGGACCGACAACATTTCTACCATCCTTGA	2074
Dd	4930	ATCAGGCAAAATGGCGTAATATTTCGTGTGCAAGGACCGACAACATTTCTACCATCCTTGA	4989
Qy	2075	CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGGTCTGCAATCATCGCATCTGTTATA	2134
Dd	4990	CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGGTCTGCAATCATCGCATCTGTTATA	5049
Qy	2135	CGACGAGCGTTCCGTCCTTHAATGTAAGACGAGTTAAGCAATCAGATCTTCTTCAGGTTATG	2194
Dd	5050	CGACGAGCGTTCCGTCCTTHAATGTAAGACGAGTTAAGCAATCAGATCTTCTTCAGGTTATG	5109
Qy	2195	ACCATCTGTGCCAGTTTCGTAATGCTGTGGTCAACTTTCCGACTCTGAGAAACTTCTGGAAT	2254
Dd	5110	ACCATCTGTGCCAGTTTCGTAATGCTGTGGTCAACTTTCCGACTCTGAGAAACTTCTGGAAT	5169
Qy	2255	CGCTAGAGAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACCGGATATATAGTGG	2314
Dd	5170	CGCTAGAGAATTTCTGNAATGGGATTCAGGAGTGGACAGAACGACCGGATATATAGTGG	5229
Qy	2315	ATGTGTCAAAACCGCATACCAATTTTGAA	2341
Dd	5230	ATGTGTCAAAACCGCATACCAATTTTGAA	5256

## RESULT 6

US-09-859-855A-2

; Sequence 2, Application US/09869855A  
; Publication No. US20030044940A1  
; GENERAL INFORMATION:  
; APPLICANT: Rasmussen, Michael Dolberg  
; TITLE OF INVENTION: Method For Increasing Gene Copy Number  
; FILE REFERENCE: 10028.204-US  
; CURRENT APPLICATION NUMBER: US/09/869.855A  
; CURRENT FILING DATE: 2001-07-05  
; NUMBER OF SEQ ID NOS: 12  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 5943  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic

US-09-859-855A-2

Query Match 27.5%; Score 867; DB 3; Length 5943;  
Best Local Similarity 100.0%; Pred. No. 7.5e-207;  
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1475	TGCTTAGGAACGAGTTATTAAAGCTGTAATAAGAACGGTGCTCTCCAAATATTTCTAT	1534
Db	968	TGCTTAGGAACGAGTTATTAAAGCTGTAATAAGAACGGTGCTCTCCAAATATTTCTAT	1027
Qy	1535	TTAGAAAAGCAAAATCTAAATTTATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA	1594
Db	1028	TTAGAAAAGCAAAATCTAAATTTATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA	1087
Qy	1595	TAATGACTAGAGAAGAAGATGAAGATTGTCATGAAATTAAGAAACGAATATTGGATA	1654
Db	1088	TAATGACTAGAGAAGAAGATGAAGATTGTCATGAAATTAAGAAACGAATATTGGATA	1147
Qy	1655	AATATGGGATGATGTTAAGGCTATTGGTGTTTATGGCTCTTGGTCTGAGACTGATG	1714
Db	1148	AATATGGGATGATGTTAAGGCTATTGGTGTTTATGGCTCTTGGTCTGAGACTGATG	1207
Qy	1715	GGCCCTATTTCGGATATTGAGATGATGTTGTCATGTCACAGAGAAAGCAGAGTTTCAGCC	1774
Db	1208	GGCCCTATTTCGGATATTGAGATGATGTTGTCATGTCACAGAGAAAGCAGAGTTTCAGCC	1267
Qy	1775	ATGAATGACAAACCGTGAGTGGAAAGTGGAAAGTGAATTTTGTAGCGAAGAGATTTCTAC	1834
Db	1368	ATGAATGACAAACCGTGAGTGGAAAGTGGAAAGTGAATTTTGTAGCGAAGAGATTTCTAC	1327
Qy	1835	TAGATTATGCAATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA	1894
Db	1328	TAGATTATGCAATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA	1387
Qy	1895	TTTTGCCGATTTATGATTTCAGGTGGATACTATTAGAGAAAGTGATCAAACTGCTAAATCGG	1954
Db	1388	TTTTGCCGATTTATGATTTCAGGTGGATACTATTAGAGAAAGTGATCAAACTGCTAAATCGG	1447
Qy	1955	TAGAAGCCAAACGGTTCCAGATGGATTGTCGCTTATCGTAGAGAGCTGTTGAAAT	2014
Db	1448	TAGAAGCCAAACGGTTCCAGATGGATTGTCGCTTATCGTAGAGAGCTGTTGAAAT	1507
Qy	2015	ATGCAGGCAAAATGGGTAATATTCGTGCAAGGACCGACAAATTTCTACCATCCTTGA	2074
Db	1508	ATGCAGGCAAAATGGGTAATATTCGTGCAAGGACCGACAAATTTCTACCATCCTTGA	1567
Qy	2075	CTGTACAGGTAGCAATGGCAGGTGCAATGTTGATTGGTCTGCATCATCGCATCTGTTATA	2134
Db	1568	CTGTACAGGTAGCAATGGCAGGTGCAATGTTGATTGGTCTGCATCATCGCATCTGTTATA	1627
Qy	2135	CGACGAGCGCTTCGGTCTTAATCTAGCAGATTAGCAATCAGATCTTCCTTCAGGTTATG	2194
Db	1628	CGACGAGCGCTTCGGTCTTAATCTAGCAGATTAGCAATCAGATCTTCCTTCAGGTTATG	1687

Qy	2195	ACCATCTGTGCCAGTTCTGTAATGTCTGTAATGCTCCGACTCTGAGAAACTTCTGGAAT	2254
Db	1688	ACCATCTGTGCCAGTTCTGTAATGTCTGTAATGCTCCGACTCTGAGAAACTTCTGGAAT	1747
Qy	2255	CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACAGGATATATAGTGG	2314
Db	1748	CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACAGGATATATAGTGG	1807
Qy	2315	ATGTGTCAAAACGCAATACCATTTTGAA	2341
Db	1808	ATGTGTCAAAACGCAATACCATTTTGAA	1834

## RESULT 7

US-10-402-466A-1

; Sequence 1, Application US/10402466A  
; Publication No. US20040028695A1  
; GENERAL INFORMATION:  
; APPLICANT: Park, Sukjoon  
; APPLICANT: Giri, Lallan  
; TITLE OF INVENTION: RECOMBINANT IMMUNOGENIC COMPOSITIONS AND METHODS FOR PROTECTING  
; TITLE OF INVENTION: AGAINST LETHAL INFECTIONS FROM BACILLUS ANTHRACIS  
; FILE REFERENCE: 18933-00005  
; CURRENT APPLICATION NUMBER: US/10/402.466A  
; CURRENT FILING DATE: 2003-03-20  
; PRIOR APPLICATION NUMBER: 60/372,152  
; PRIOR FILING DATE: 2002-04-12  
; NUMBER OF SEQ ID NOS: 34  
; SEQ ID NO 1  
; LENGTH: 6694  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; OTHER INFORMATION: Artificial DNA sequence to be used as one of the two backbones of  
; OTHER INFORMATION: E. coli-Bacillus anthracis shuttle vectors. Designated as pBP 1.  
US-10-402-466A-1

Query Match 27.5%; Score 867; DB 7; Length 6694;  
Best Local Similarity 100.0%; Pred. No. 8e-207;  
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1475	TGCTTAGGAACGAGTTATTAAATAGCTGAATAAGAACGGTGCTCTCCAAATATTTCTAT	1534
Db	5219	TGCTTAGGAACGAGTTATTAAATAGCTGAATAAGAACGGTGCTCTCCAAATATTTCTAT	5278
Qy	1535	TTAGAAAAGCAAAATCTAAATTTATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA	1594
Db	5279	TTAGAAAAGCAAAATCTAAATTTATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA	5338
Qy	1595	TAATGACTAGAGAAGAAGATGAAGATTGTCATGAAATTAAGAAACGAATATTGGATA	1654
Db	5339	TAATGACTAGAGAAGAAGATGAAGATTGTCATGAAATTAAGAAACGAATATTGGATA	5398
Qy	1655	AATATGGGATGATGTTAAGGCTATTGCTGTTTATGGCTCTCTTGGTCTGAGACTGATG	1714
Db	5399	AATATGGGATGATGTTAAGGCTATTGCTGTTTATGGCTCTCTTGGTCTGAGACTGATG	5458
Qy	1715	GGCCCTATTTCGGATATTGAGATGATGTTGTCATGTCAACAGAGAAAGCAGAGTTTCAGCC	1774
Db	5459	GGCCCTATTTCGGATATTGAGATGATGTTGTCATGTCAACAGAGAAAGCAGAGTTTCAGCC	5518
Qy	1775	ATGAATGGACAAACCGGTGAGTGGAAAGTGAATTTTGTAGAGCAAGAGATTTCTAC	1834
Db	5519	ATGAATGGACAAACCGGTGAGTGGAAAGTGAATTTTGTAGAGCAAGAGATTTCTAC	5578
Qy	1835	TAGATTATGCAATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA	1894
Db	5579	TAGATTATGCAATCTCAGGTGGAATCAGATTGGCCGCTTACACATGGTCAATTTTCTCTA	5638
Qy	1895	TTTTGCCGATTTATGATTTCAGGTGGATACTATTAGAGAAAGTGATCAAACTGCTAAATCGG	1954

Db 5639 TTTTGGCGATTTATGATTAGGTGGATACCTAGAGAAAGTGTATCAAACTGCTAAATCGG 5698  
Qy 1955 TAGAAGCCCAAACTTCCAGATGCGATTTGTGCCCTTATCGTAGAAGAGCTGTTTGAAT 2014  
Db 5699 TAGAAGCCCAAACTTCCAGATGCGATTTGTGCCCTTATCGTAGAAGAGCTGTTTGAAT 5758  
Qy 2015 ATGCAGGCAAAATGCGGTAATTTTGTGTGCAAGGACCGACAACATTTCTACCATCTCTGA 2074  
Db 5759 ATGCAGGCAAAATGCGGTAATTTTGTGTGCAAGGACCGACAACATTTCTACCATCTCTGA 5818  
Qy 2075 CTGTACAGGTAGCAATGCGAGGTGCCATGTTGTTGCTTGCATCATCGCATCTGTTATA 2134  
Db 5819 CTGTACAGGTAGCAATGCGAGGTGCCATGTTGTTGCTTGCATCATCGCATCTGTTATA 5878  
Qy 2135 CGACGAGCGCTTCGGTCTTAACGAAGCAGTTAAGCAATCAGATCTTCTTCCAGGTTATG 2194  
Db 5879 CGACGAGCGCTTCGGTCTTAACGAAGCAGTTAAGCAATCAGATCTTCTTCCAGGTTATG 5938  
Qy 2195 ACCATCTGTGCCAGTTCGTAAATGTTCTGCTCAACTTTCCGACTCTGAGAAAATTTCTGGAAT 2254  
Db 5939 ACCATCTGTGCCAGTTCGTAAATGTTCTGCTCAACTTTCCGACTCTGAGAAAATTTCTGGAAT 5998  
Qy 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGAGTGGACAGAACGACGAGATATATAGTGG 2314  
Db 5999 CGCTAGAGAAATTTCTGGAATGGGATTCAGAGTGGACAGAACGACGAGATATATAGTGG 6058  
Qy 2315 ATGTGTCAAAACGCATACCATTTTGA 2341  
Db 6059 ATGTGTCAAAACGCATACCATTTTGA 6085

RESULT 8

US-10-084-814-3  
; Sequence 3, Application US/10084814  
; Publication No. US20030108982A1  
; GENERAL INFORMATION:  
; APPLICANT: SLIJKHUIS, HERMAN; SELTEN,  
; ERIC BASTIAAN  
; TITLE OF INVENTION: PROCESS FOR OXIDATION OF  
; STEROIDS AND GENETICALLY ENGINEERED CELLS  
; USED THEREIN  
; NUMBER OF SEQUENCES: 79  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BIERMAN, MUSERLIAN & LUCAS  
; STREET: 600 THIRD AVENUE  
; CITY: NEW YORK  
; STATE: NEW YORK  
; COUNTRY: USA  
; ZIP: 10016  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: FLOPPY DISK  
; COMPUTER: IBM PC COMPATIBLE  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: MICROSOFT WORD 97  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/10/084,814  
; FILING DATE: 26-Feb-2002  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/418,085  
; FILING DATE: 06-APR-1995  
; APPLICATION NUMBER: 08/054,185  
; FILING DATE: 26-APR-1993  
; APPLICATION NUMBER: 08/002,608  
; FILING DATE: 11-JAN-1993  
; APPLICATION NUMBER: 07/474,857  
; FILING DATE: 30-OCT-1990  
; APPLICATION NUMBER: 07/474,798  
; FILING DATE: 16-JULY-1990  
; APPLICATION NUMBER: PCT/NL89/00072  
; FILING DATE: 25-SEPT-1989  
; APPLICATION NUMBER: NL88/200904.6

; FILING DATE: 06-MAY-1988  
; APPLICATION NUMBER: NL/88/202080.3  
; FILING DATE: 03-SEP-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: CHARLES A. MUSERLIAN  
; REGISTRATION NUMBER: 19,683  
; REFERENCE/DOCKET NUMBER: 146,1169-  
; CON-1-DIV-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 661-8000  
; TELEFAX: (212) 661-8002  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7336 BASE PAIRS  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: DOUBLE  
; TOPOLOGY: UNKNOWN  
; FEATURE:  
; OTHER INFORMATION: PLASMID pBHA-1  
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-10-084-814-3  
Query Match 27.5%; Score 867; DB 5; Length 7336;  
Best Local Similarity 100.0%; Pred. No. 8.4e-207; Indels 0; Gaps 0;  
Matches 867; Conservative 0; Mismatches 0;  
Qy 1475 TGCTTAGGAGACAGAGTTAATAAGCTGAATGAAGACGGTCTCTCCAAATATTCCTAT 1534  
Db 5231 TGCTTAGGAGACAGAGTTAATAAGCTGAATGAAGACGGTCTCTCCAAATATTCCTAT 5290  
Qy 1535 TTAGAAAAGCAAAATCTAAAATTTATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA 1594  
Db 5291 TTAGAAAAGCAAAATCTAAAATTTATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA 5350  
Qy 1595 TAATGACTAGAGAAGAAAGAAATCAAGATTTGCTCATGAAATTAAGGAACGAATATTGATA 1654  
Db 5351 TAATGACTAGAGAAGAAAGAAATCAAGATTTGCTCATGAAATTAAGGAACGAATATTGATA 5410  
Qy 1655 AATATGGGATGATGTTAAGCTTATTTGCTTTATGGCTCTCTTTGGTCTGAGACTGATG 1714  
Db 5411 AATATGGGATGATGTTAAGCTTATTTGCTTTATGGCTCTCTTTGGTCTGAGACTGATG 5470  
Qy 1715 GGCCCTATTCCGATATTGAGATGATGTTGCTCATGTCATGCAACAGAGGAAGAGAGTTACGCC 1774  
Db 5471 GGCCCTATTCCGATATTGAGATGATGTTGCTCATGTCATGTCACAGAGGAAGAGAGTTACGCC 5530  
Qy 1775 ATGAATGGCAACCGGTGAGTGAAGTGAATTTTGTAGTAGGAGAGAGATTCTAC 1834  
Db 5531 ATGAATGGCAACCGGTGAGTGAAGTGAATTTTGTAGTAGGAGAGAGATTCTAC 5590  
Qy 1835 TAGATTATGATCTCAGGTGGAATCAGATTGGCGGCTTACACATGTTGTCATTTTCTCTA 1894  
Db 5591 TAGATTATGATCTCAGGTGGAATCAGATTGGCGGCTTACACATGTTGTCATTTTCTCTA 5650  
Qy 1895 TTTTGGCGATTTATGATTTAGTGGATATTAGAGAAAGTGTATCAAACTGCTAAATCGG 1954  
Db 5651 TTTTGGCGATTTATGATTTAGTGGATATTAGAGAAAGTGTATCAAACTGCTAAATCGG 5710  
Qy 1955 TAGAAGCCCAAAAGTTTCCAGATGCGATTTGTGCCCTTATCGTAGAAGAGCTGTTGAAT 2014  
Db 5711 TAGAAGCCCAAAAGTTTCCAGATGCGATTTGTGCCCTTATCGTAGAAGAGCTGTTGAAT 5770  
Qy 2015 ATGCAGGCAAAATGCGGTAATTTTGTGTGCAAGGACCGACAACATTTCTACCATCTCTGA 2074  
Db 5771 ATGCAGGCAAAATGCGGTAATTTTGTGTGCAAGGACCGACAACATTTCTACCATCTCTGA 5830  
Qy 2075 CTGTACAGGTAGCAATGCGAGGTGCCATGTTGTTGCTTGCATCATCGCATCTGTTATA 2134  
Db 5831 CTGTACAGGTAGCAATGCGAGGTGCCATGTTGTTGCTTGCATCATCGCATCTGTTATA 5890  
Qy 2135 CGACGAGCGCTTCGGTCTTAACGAAGCAGTTAAGCAATCAGATCTTCTTCCAGGTTATG 2194  
Db 5891 CGACGAGCGCTTCGGTCTTAACGAAGCAGTTAAGCAATCAGATCTTCTTCCAGGTTATG 5950

Qy	2195	ACCATCTGTGCCAGTTTCGTAAATGCTCGTCAACTTTCCGACTCTGAGAAA	CTTCTCGGAAT	2254
Db	5951	ACCATCTGTGCCAGTTTCGTAAATGCTCGTCAACTTTCCGACTCTGAGAAA	CTTCTCGGAAT	6010
Qy	2255	CGCTAGAGAATTTCTCGAATGGGATTTCAGGAGTGGACAGAACGACCGGATATATAGTGG	2314	
Db	6011	CGCTAGAGAATTTCTCGAATGGGATTTCAGGAGTGGACAGAACGACCGGATATATAGTGG	2314	
Qy	2315	ATGTGTCAAAACGCATACCATTTTGAA	2341	
Db	6071	ATGTGTCAAAACGCATACCATTTTGAA	6097	

```

RESULT 9
US-10-442-013-4
; Sequence 4, Application US/10442013
; Publication No. US20040014175A1
; GENERAL INFORMATION:
; APPLICANT: van der Laan, Johannes Cornelis
; APPLICANT: van Eekelen, Christiaan Albertus Gerardus
; TITLE OF INVENTION: Efficient Production of Mutant Proteases
; FILE REFERENCE: GCX329-US-C1
; CURRENT APPLICATION NUMBER: US/10/442,013
; CURRENT FILING DATE: 2003-05-19
; PRIOR APPLICATION NUMBER: EP 89202117.1
; PRIOR FILING DATE: 1989-08-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 7336
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: plasmid pBHA-1
US-10-442-013-4

```

Qy	1955	TAGAAGCCCAAAACGTTTCCAGATGCGATTTGGCCCTTATCGTAGAAGAGCTGTTTGAAT	2011
Db	5711	TAGAAGCCCAAAACGTTTCCAGATGCGATTTGGCCCTTATCGTAGAAGAGCTGTTTGAAT	5770
Qy	2015	ATCAGCGCAAAATCGCGTAAATATTCGTGTCNAGGACCGACAACATTTCTACCATCCTTGA	2074
Db	5771	ATCAGCGCAAAATCGCGTAAATATTCGTGTCNAGGACCGACAACATTTCTACCATCCTTGA	5830
Qy	2075	CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGGTCTGTCATCATCGCATCTGTTATA	2134
Db	5831	CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGGTCTGTCATCATCGCATCTGTTATA	5890
Qy	2135	CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAAGCAATCAGATCTTCTTCAGGTTATG	2194
Db	5891	CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAAGCAATCAGATCTTCTTCAGGTTATG	5950
Qy	2195	ACCATCTGTGCGCAGTTTCGTAATGTCGTGTCCTCAACTTTCCGACTCTGAGAACTTCTGGAAT	2254
Db	5951	ACCATCTGTGCGCAGTTTCGTAATGTCGTGTCCTCAACTTTCCGACTCTGAGAACTTCTGGAAT	6010
Qy	2255	CGCTAGAGAAATTTCTGGAATGGATTCAGAGTGGACAGAAACGACACGGATATATAGTGG	2314
Db	6011	CGCTAGAGAAATTTCTGGAATGGATTCAGAGTGGACAGAAACGACACGGATATATAGTGG	6070
Qy	2315	ATGTGTCAAAACGCATACCATTTTGAA	2341
Db	6071	ATGTGTCAAAACGCATACCATTTTGAA	6097
RESULT 10			
US-10-462-128-3			
; Sequence 3, Application US/10462128			
; Publication No. US20040067579A1			
; GENERAL INFORMATION:			
; TITLE OF INVENTION: PROCESS FOR OXIDATION OF STEROIDS AND			
; GENETICALLY ENGINEERED CELLS USED THEREIN			
; NUMBER OF SEQUENCES: 79			
; CORRESPONDENCE ADDRESS:			
; ADDRESSEE: BIERMAN & MUSERLIAN			
; STREET: 600 THIRD AVENUE			
; CITY: NEW YORK			
; STATE: NEW YORK			
; COUNTRY: USA			
; ZIP: 10016			
; COMPUTER READABLE FORM:			
; MEDIUM TYPE: FLOPPY DISK			
; COMPUTER: IBM PC COMPATIBLE			
; OPERATING SYSTEM: PC-DOS/MS-DOS			
; SOFTWARE: ASCII			
; CURRENT APPLICATION DATA:			
; APPLICATION NUMBER: US/10/462,128			
; FILING DATE: 16-Jun-2003			
; CLASSIFICATION: <Unknown>			
; PRIOR APPLICATION DATA:			
; APPLICATION NUMBER: US/08/054,185			
; FILING DATE: 26-APR-1993			
; APPLICATION NUMBER: US/08/002,608			
; FILING DATE: 11-JAN-1993			
; APPLICATION NUMBER: US/07/474,857			
; FILING DATE: 30-OCT-1990			
; APPLICATION NUMBER: US/07/474,798			
; FILING DATE: 16-JULY-1990			
; APPLICATION NUMBER: PCT/NL89/00072			
; FILING DATE: 25-SEPT-1989			
; APPLICATION NUMBER: NL/88/200904.6			
; FILING DATE: 06-MAY-88			
; APPLICATION NUMBER: NL/88/202080.3			
; FILING DATE: 03-SEP-88			
; ATTORNEY/AGENT INFORMATION:			
; NAME: CHARLES A. MUSERLIAN			
; REGISTRATION NUMBER: 19,683			
; REFERENCE/DOCKET NUMBER: 146.1169 CON-1			
; TELECOMMUNICATION INFORMATION:			

TELEPHONE: (212) 661-8000  
TELEFAX: (212) 661-8002  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 7336  
TYPE: NUCLEIC ACID  
STRANDEDNESS: DOUBLE  
TOPOLOGY: UNKNOWN  
FEATURE:  
NAME/KEY:  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: PLASMID pBHA-1  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-10-462-128-3

Query Match 27.5%; Score 867; DB 7; Length 7336;  
Best Local Similarity 100.0%; Pred. No. 8.4e-207;  
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1475 TGCCTAGGAAGACGAGTTAATAAGTGAATGAAGACGGTCTCTCCAAATATTCCTAT 1534  
DB 5231 TGCCTAGGAAGACGAGTTAATAAGTGAATGAAGACGGTCTCTCCAAATATTCCTAT 5290  
QY 1535 TTAGAAAGCAATCTAAATATCTGAAAGGAATGAGAATAGTGAATGGACCAATAA 1594  
DB 5291 TTAGAAAGCAATCTAAATATCTGAAAGGAATGAGAATAGTGAATGGACCAATAA 5350  
QY 1595 TAATGACTAGAGAAGAAATGAAGATGCTTCAATGAATTAAGGAACGAATATTCGATA 1654  
DB 5351 TAATGACTAGAGAAGAAATGAAGATGCTTCAATGAATTAAGGAACGAATATTCGATA 5410  
QY 1655 AATATGGGATGATGTTAAGGCTATTGCTGTTTATGGCTCTCTTGGTCTGTCAGACTGATG 1714  
DB 5411 AATATGGGATGATGTTAAGGCTATTGCTGTTTATGGCTCTCTTGGTCTGTCAGACTGATG 5470  
QY 1715 GGCCCTATTCGGATATTGAGATGATGTCATGTCAATGCAACAGAGGAAGCAGAGTTCAGCC 1774  
DB 5471 GGCCCTATTCGGATATTGAGATGATGTCATGTCAATGCAACAGAGGAAGCAGAGTTCAGCC 5530  
QY 1775 ATGAATGGACAACCGGTGAGTGAAGTGAAGTGAATTTTGTATAGGAAGAGATTCCTAC 1834  
DB 5531 ATGAATGGACAACCGGTGAGTGAAGTGAAGTGAATTTTGTATAGGAAGAGATTCCTAC 5590  
QY 1835 TAGATTATGATCTCAGTGGAAATCAGATGGCGCTTACACATGGTCAATTTTCTCTA 1894  
DB 5591 TAGATTATGATCTCAGTGGAAATCAGATGGCGCTTACACATGGTCAATTTTCTCTA 5650  
QY 1895 TTTTGGCGATTTATGATTCAGGTGGATACCTTAGAGAAGTGTATCAAACTGCTAAATCGG 1954  
DB 5651 TTTTGGCGATTTATGATTCAGGTGGATACCTTAGAGAAGTGTATCAAACTGCTAAATCGG 5710  
QY 1955 TAGAAGCCAAACGTTCCAGATGCGATTTTGGCCCTTATCGTAGAAGAGCTGTTGAAT 2014  
DB 5711 TAGAAGCCAAACGTTCCAGATGCGATTTTGGCCCTTATCGTAGAAGAGCTGTTGAAT 5770  
QY 2015 ATCAGGCAATCGCGTAAATATTCGTGTCAAGGACGCAACATTTCTACCACTCTTGA 2074  
DB 5771 ATCAGGCAATCGCGTAAATATTCGTGTCAAGGACGCAACATTTCTACCACTCTTGA 5830  
QY 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTTGCTCTGCATCATCGATCTGTTATA 2134  
DB 5831 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTTGCTCTGCATCATCGATCTGTTATA 5890  
QY 2135 CGACGAGCGTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 2194  
DB 5891 CGACGAGCGTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 5950  
QY 2195 ACCATCTGTCAGTTCGTAATGTCGTCAACTTTTCCGACTCTGAGAACTCTCTGGAAT 2254  
DB 5951 ACCATCTGTCAGTTCGTAATGTCGTCAACTTTTCCGACTCTGAGAACTCTCTGGAAT 6010  
QY 2255 CGCTAGAAATTTCTGGAATGGGATTTCAGGAGTGGACAGAACGACGGATATATAGTGG 2314

DB 6011 CGCTAGAGAATTTCTGGAATGGGATTCAGAGTGGACAGAACGACCGATATATAGTGG 6070  
QY 2315 ATGTGTCAAAACGACATACCATTTTGAA 2341  
DB 6071 ATGTGTCAAAACGACATACCATTTTGAA 6097  
RESULT 11  
US-10-402-466A-7  
; Sequence 7, Application US/10402466A  
; Publication No. US20040028695A1  
; GENERAL INFORMATION:  
; APPLICANT: Park, Sukjoon  
; APPLICANT: Giri, Lallan  
; TITLE OF INVENTION: RECOMBINANT IMMUNOGENIC COMPOSITIONS AND METHODS FOR PROTECTING  
; FILE REFERENCE: 18933-00005  
; CURRENT APPLICATION NUMBER: US/10/402,466A  
; PRIOR FILING DATE: 2003-03-20  
; PRIOR FILING DATE: 2002-04-12  
; NUMBER OF SEQ ID NOS: 34  
; SEQ ID NO 7  
; LENGTH: 8198  
; TYPE: DNA  
; ORGANISM: Bacillus anthracis  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; OTHER INFORMATION: Expression vector for B. anthracis Protective Antigen (PA). Enti  
; OTHER INFORMATION: sequence is shown since the vector sequence is different from the  
; OTHER INFORMATION: of the pBP vectors. The PA coding sequence is from 3735 to 5942.  
US-10-402-466A-7  
Query Match 27.5%; Score 867; DB 7; Length 8198;  
Best Local Similarity 100.0%; Pred. No. 8.9e-207;  
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1475 TGCCTAGGAAGACGAGTTAATAAGTGAATGAAGACGGTCTCTCCAAATATTCCTAT 1534  
DB 2259 TGCCTAGGAAGACGAGTTAATAAGTGAATGAAGACGGTCTCTCCAAATATTCCTAT 2318  
QY 1535 TTAGAAAGCAATCTAAATATCTGAAAGGAATGAGAATAGTGAATGGACCAATAA 1594  
DB 2319 TTAGAAAGCAATCTAAATATCTGAAAGGAATGAGAATAGTGAATGGACCAATAA 2378  
QY 1595 TAATGACTAGAGAAGAAATGAAGATGCTTCAATGAATTAAGGAACGAATATTCGATA 1654  
DB 2379 TAATGACTAGAGAAGAAATGAAGATGCTTCAATGAATTAAGGAACGAATATTCGATA 2438  
QY 1655 AATATGGGATGATGTTAAGGCTATTGCTGTTTATGGCTCTCTTGGTCTGTCAGACTGATG 1714  
DB 2439 AATATGGGATGATGTTAAGGCTATTGCTGTTTATGGCTCTCTTGGTCTGTCAGACTGATG 2498  
QY 1715 GGCCCTATTCGGATATTGAGATGATGTCATGTCAACAGAGGAAGCAGAGTTCAGCC 1774  
DB 2499 GGCCCTATTCGGATATTGAGATGATGTCATGTCAACAGAGGAAGCAGAGTTCAGCC 2558  
QY 1775 ATGAATGGACAACCGGTGAGTGAAGTGAATTTTGTATAGGAAGAGATTCCTAC 1834  
DB 2559 ATGAATGGACAACCGGTGAGTGAAGTGAATTTTGTATAGGAAGAGATTCCTAC 2618  
QY 1835 TAGATTATGATCTCAGTGGAAATCAGATGGCGCTTACACATGGTCAATTTTCTCTA 1894  
DB 2619 TAGATTATGATCTCAGTGGAAATCAGATGGCGCTTACACATGGTCAATTTTCTCTA 2678  
QY 1895 TTTTGGCGATTTATGATTCAGGTGGATACCTTAGAGAAGTGTATCAAACTGCTAAATCGG 1954  
DB 2679 TTTTGGCGATTTATGATTCAGGTGGATACCTTAGAGAAGTGTATCAAACTGCTAAATCGG 2738  
QY 1955 TAGAAGCCAAACGTTCCAGATGCGATTTTGGCCCTTATCGTAGAAGAGCTGTTGAAT 2014



Db 2739 TAGAAGCCAAACGTTCCACGATGCGATTTGTGCGCCTTATCGTAGAGAGCTGTTTGAAT 2798  
Qy 2015 ATGCAGGCAANTGGCGTAATATTTGTTGCGCAAGACCGACAAACATTTCTACCATCCTTGA 2074  
Db 2799 ATGCAGGCAANTGGCGTAATATTTGTTGCGCAAGACCGACAAACATTTCTACCATCCTTGA 2858  
Qy 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGTTGCTGCATCATCGCATCTGTTTATA 2134  
Db 2859 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGTTGCTGCATCATCGCATCTGTTTATA 2918  
Qy 2135 CGACAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 2194  
Db 2919 CGACAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 2978  
Qy 2195 ACCATCTGTGCCAGTTCGTATATGCTGTCGTAACCTTCGCACTCTGAGNAACCTTCGGAAT 2254  
Db 2979 ACCATCTGTGCCAGTTCGTATATGCTGTCGTAACCTTCGCACTCTGAGNAACCTTCGGAAT 3038  
Qy 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGACGACACGGATATATAGTGG 2314  
Db 3039 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGACGACACGGATATATAGTGG 3098  
Qy 2315 ATGTGTCAAAACGCATACCATTTTGAA 2341  
Db 3099 ATGTGTCAAAACGCATACCATTTTGAA 3125

RESULT 12  
US-10-402-466A-10  
; Sequence 10, Application US/10402466A  
; Publication No. US20040028695A1  
; GENERAL INFORMATION:  
; APPLICANT: Park, Sukjoon  
; APPLICANT: Giri, Laljan  
; TITLE OF INVENTION: RECOMBINANT IMMUNOGENIC COMPOSITIONS AND METHODS FOR PROTECTING  
; TITLE OF INVENTION: AGAINST LETHAL INFECTIONS FROM BACILLUS ANTHRACIS  
; FILE REFERENCE: 18933-00005  
; CURRENT APPLICATION NUMBER: US/10/402,466A  
; CURRENT FILING DATE: 2003-03-20  
; PRIOR APPLICATION NUMBER: 60/372,152  
; PRIOR FILING DATE: 2002-04-12  
; NUMBER OF SEQ ID NOS: 34  
; SEQ ID NO 10  
; LENGTH: 9286  
; TYPE: DNA  
; ORGANISM: Bacillus anthracis  
; FEATURE:  
; NAME/KEY:  
; LOCATION:  
; OTHER INFORMATION: Dual expression vector pBP105 for B. anthracis PA and LF30. Enti  
; OTHER INFORMATION: sequence is shown since the vector sequence contains two coding  
; OTHER INFORMATION: regions. The coding region for PA is from 3735 to 5942 and the c  
; OTHER INFORMATION: region for LF30 is from 6391 to 7161.  
US-10-402-466A-10

Query Match 27.5%; Score 867; DB 7; Length 9286;  
Best Local Similarity 100.0%; Pred. No. 9.6e-207; Mismatches 0; Indels 0; Gaps 0;  
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1475 TGCTTAGGAACGAGTTTATTAATAGCTGAATAAGAACGGTGCCTCCAAATATTTCTTAT 1534  
Db 2259 TGCTTAGGAACGAGTTTATTAATAGCTGAATAAGAACGGTGCCTCCAAATATTTCTTAT 2318  
Qy 1535 TTAGAAAAGCAATCTAAATATTTCTGAAAAGGGAATGAGATAGTGAATGAGCAATATA 1594  
Db 2319 TTAGAAAAGCAATCTAAATATTTCTGAAAAGGGAATGAGATAGTGAATGAGCAATATA 2378  
Qy 1595 TAATGACTAGAGAGAAGAGATGAGATTTGTCATGAATTAAGAAACGATATTTGGATA 1654  
Db 2379 TAATGACTAGAGAGAAGAGATGAGATTTGTCATGAATTAAGAAACGATATTTGGATA 2438  
Qy 1655 AATATGGGATGATGTTAAGGCTATTTGGTGTATGCTCTCTTGGTGTGCTCAGACTGATG 1714

Db 2439 AATATGGGATGATGTTAAGGCTATTTGGTGTATGCTCTCTTGGTGTGCTCAGACTGATG 2498  
Qy 1715 GGCCTTATTCGGATATTGAGATGATGTTGCTCATGTCTCAATGTCACAGAGGAGACAGATTTCAGCC 1774  
Db 2499 GGCCTTATTCGGATATTGAGATGATGTTGCTCATGTCTCAATGTCACAGAGGAGACAGATTTCAGCC 2558  
Qy 1775 ATGAATGGACAAACCGGTGAGTGGAAAGGTGGAAGTGAATTTTGTAGCGAAGAGATTCTTAC 1834  
Db 2559 ATGAATGGACAAACCGGTGAGTGGAAAGGTGGAAGTGAATTTTGTAGCGAAGAGATTCTTAC 2618  
Qy 1835 TAGATTAATGATCTCAGGTGGAATCAGATTTGGCGCTTACACATGTTCAATTTTCTCTA 1894  
Db 2619 TAGATTAATGATCTCAGGTGGAATCAGATTTGGCGCTTACACATGTTCAATTTTCTCTA 2678  
Qy 1895 TTTTGGCGGATTTATGATTTAGGTTGGATTTAGGAAAGTGTATCAAACTGCTAAATCGG 1954  
Db 2679 TTTTGGCGGATTTATGATTTAGGTTGGATTTAGGAAAGTGTATCAAACTGCTAAATCGG 2738  
Qy 1955 TAGAAGCCCAAAACGTTCCACGATGCGATTTGTGCTTATCGTAGAAGAGAGCTGTTTGAAT 2014  
Db 2739 TAGAAGCCCAAAACGTTCCACGATGCGATTTGTGCTTATCGTAGAAGAGAGCTGTTTGAAT 2798  
Qy 2015 ATGCAGGCAATGCGCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 2074  
Db 2799 ATGCAGGCAATGCGCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 2858  
Qy 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGCTGTCATCGCATCTGTTTATA 2134  
Db 2859 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGCTGTCATCGCATCTGTTTATA 2918  
Qy 2135 CGACAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 2194  
Db 2919 CGACAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 2978  
Qy 2195 ACCATCTGTGCCAGTTCGTATATGCTGTTGCTCAACTTCGCGACTGTGAGAAACCTTCGGAAT 2254  
Db 2979 ACCATCTGTGCCAGTTCGTATATGCTGTTGCTCAACTTCGCGACTGTGAGAAACCTTCGGAAT 3038  
Qy 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGACGACACGGATATATAGTGG 2314  
Db 3039 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGACGACACGGATATATAGTGG 3098  
Qy 2315 ATGTGTCAAAACGCATACCATTTTGAA 2341  
Db 3099 ATGTGTCAAAACGCATACCATTTTGAA 3125

RESULT 13  
US-10-032-393-7/c  
; Sequence 7, Application US/10032393  
; Publication No. US20030027286A1  
; GENERAL INFORMATION:  
; APPLICANT: Haeelbeck, Robert  
; APPLICANT: Wall, Daniel  
; APPLICANT: Gross, Molly  
; TITLE OF INVENTION: BACTERIAL PROMOTERS AND METHODS OF USE  
; FILE REFERENCE: ELITEA 010A  
; CURRENT APPLICATION NUMBER: US/10/032,393  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/259,434  
; PRIOR FILING DATE: 2000-12-27  
; PRIOR APPLICATION NUMBER: 09/948,993  
; PRIOR FILING DATE: 2001-09-06  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; NUMBER OF SEQ ID NOS: 68  
; SOFTWARE: Fast-Seq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 10929  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Vector pBP25



US-10-032-393-7

Query Match 27.5%; Score 867; DB 5; Length 10929;  
Best Local Similarity 100.0%; Pred. No. 1e-206;  
Matches 867; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1475 TGCTTAGGAAGACGAGTTATTAATAGCTGAATAGAACGGTGTCTCTCAAAATATCTTAT 1534  
DB 8753 TGCTTAGGAAGACGAGTTATTAATAGCTGAATAGAACGGTGTCTCTCAAAATATCTTAT 8694

QY 1535 TTAGAAAAGCAATCTAAATATCTGAAAAGGAATGAGATAGTCAATGGACCAATAA 1594  
DB 8693 TTAGAAAAGCAATCTAAATATCTGAAAAGGAATGAGATAGTCAATGGACCAATAA 8634

QY 1595 TAATGACTAGAGAAGAAAGTAAGATGTTTCATGAAATTAAGGAAACGAATATGGATA 1654  
DB 8633 TAATGACTAGAGAAGAAAGTAAGATGTTTCATGAAATTAAGGAAACGAATATGGATA 8574

QY 1655 AATATGGGATGATGTTAAAGCTATATGGTGTATATGCTCTCTTGGTGTGACACTGATG 1714  
DB 8573 AATATGGGATGATGTTAAAGCTATATGGTGTATATGCTCTCTTGGTGTGACACTGATG 8514

QY 1715 GGCCCTATTCGGATATTCAGATGATGTCATGTCATGTCACAGAGAACGAGTTCAGCC 1774  
DB 8513 GGCCCTATTCGGATATTCAGATGATGTCATGTCATGTCACAGAGAACGAGTTCAGCC 8454

QY 1775 ATGAATGGACAACCGGTGAGTGGAAAGTGGAAAGTGAATTTTATAGCGAAGAGATTCCTAC 1834  
DB 8453 ATGAATGGACAACCGGTGAGTGGAAAGTGGAAAGTGAATTTTATAGCGAAGAGATTCCTAC 8394

QY 1835 TAGATTAATGATCTCAGGTGGAATCAGATGGCCGCTTACACATGGTCAATTTTCTCTA 1894  
DB 8393 TAGATTAATGATCTCAGGTGGAATCAGATGGCCGCTTACACATGGTCAATTTTCTCTA 8334

QY 1895 TTTTGCCGATTTATGATTCAGTGGATACCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 1954  
DB 8333 TTTTGCCGATTTATGATTCAGTGGATACCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 8274

QY 1955 TAGAAGCCCAACCGTCCAGATCGGATTTGTGCGCTTATCGTAGAAGAGCTGTTGCAAT 2014  
DB 8273 TAGAAGCCCAACCGTCCAGATCGGATTTGTGCGCTTATCGTAGAAGAGCTGTTGCAAT 8214

QY 2015 ATGACGACAAATGGCGTAATATTCGTGTGCAAGGACCGACAACTTCTACCACTCTGA 2074  
DB 8213 ATGACGACAAATGGCGTAATATTCGTGTGCAAGGACCGACAACTTCTACCACTCTGA 8154

QY 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTTGCTGTCATCGCATCTGTTATA 2134  
DB 8153 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTTGCTGTCATCGCATCTGTTATA 8094

QY 2135 CGACGAGCGTTCGGTCTTAACTGAACAGTTAAGCAATCAGATCTTCCCTTCAGGTTATG 2194  
DB 8093 CGACGAGCGTTCGGTCTTAACTGAACAGTTAAGCAATCAGATCTTCCCTTCAGGTTATG 8034

QY 2195 ACCATCTGTGCGAGTTCGTATGTCGTGCTCAACTTTCCGACTCTGAGAACTCTCGGAAT 2254  
DB 8033 ACCATCTGTGCGAGTTCGTATGTCGTGCTCAACTTTCCGACTCTGAGAACTCTCGGAAT 7974

QY 2255 CGGTAGAGAAATTTCTGGAAATGGGATTCGGAGTGGACAGACGACAGATATAGTGG 2314  
DB 7973 CGGTAGAGAAATTTCTGGAAATGGGATTCAGGAGTGGACAGACGACAGATATAGTGG 7914

QY 2315 ATGTGTCAAAACGCATACCATTTTGAA 2341  
DB 7913 ATGTGTCAAAACGCATACCATTTTGAA 7887

RESULT 14

US-10-385-415-2/c  
; Sequence 2, Application US/10385415  
; Publication No. US20040014158A1  
; GENERAL INFORMATION:  
; APPLICANT: Bacher, Adelbert

; APPLICANT: Fischer, Markus  
; TITLE OF INVENTION: PROTEIN CONJUGATES, METHOD, VECTORS, PROTEINS AND DNA FOR  
; TITLE OF INVENTION: PRODUCING THEM, THEIR USE AND MEDICAMENTS AND VACCINES CONTAININ  
; TITLE OF INVENTION: A CERTAIN QUANTITY OF SAID PROTEIN CONJUGATES  
; FILE REFERENCE: 9286.6CT  
; CURRENT APPLICATION NUMBER: US/10/385,415  
; CURRENT FILING DATE: 2003-03-10  
; PRIOR APPLICATION NUMBER: US 09/936,028  
; PRIOR FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: PCT/EP00/01899  
; PRIOR FILING DATE: 2000-03-03  
; PRIOR APPLICATION NUMBER: DE 19910102.7  
; PRIOR FILING DATE: 1999-03-08  
; NUMBER OF SEQ ID NOS: 154  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2  
; LENGTH: 5302  
; TYPE: DNA  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: p6021-CAT Expression vector  
US-10-385-415-2

Query Match 27.4%; Score 865.4; DB 7; Length 5302;  
Best Local Similarity 99.9%; Pred. No. 1.8e-206;  
Matches 866; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1475 TGCTTAGGAAGACGAGTTATTAATAGCTGAATAGAACGGTGTCTCTCAAAATATCTTAT 1534  
DB 3821 TGCTTAGGAAGACGAGTTATTAATAGCTGAATAGAACGGTGTCTCTCAAAATATCTTAT 3762

QY 1535 TTAGAAAAGCAATCTAAATATCTGAAAAGGAATGAGAATAGTGAATGGACCAATAA 1594  
DB 3761 TTAGAAAAGCAATCTAAATATCTGAAAAGGAATGAGAATAGTGAATGGACCAATAA 3702

QY 1595 TAATGACTAGAGAAGAAAGTAAGATGTTTCATGAAATTAAGGAAACGAATATGGATA 1654  
DB 3701 TAATGACTAGAGAAGAAAGTAAGATGTTTCATGAAATTAAGGAAACGAATATGGATA 3642

QY 1655 AATATGGGATGATGTTAAAGCTATTCGTGCTTATGCTCTCTTGGTGTGTCAGACTGATG 1714  
DB 3641 AATATGGGATGATGTTAAAGCTATTCGTGCTTATGCTCTCTTGGTGTGTCAGACTGATG 3582

QY 1715 GGCCCTATTCGGATATTCAGATGATGTCATGTCATGTCACAGAGGAAGCAGAGTTCAGCC 1774  
DB 3581 GGCCCTATTCGGATATTCAGATGATGTCATGTCATGTCACAGAGGAAGCAGAGTTCAGCC 3522

QY 1775 ATGAATGGACAACCGGTGAGTGGAAAGTGAATTTTGTATAGCGAAGAGATTCCTAC 1834  
DB 3521 ATGAATGGACAACCGGTGAGTGGAAAGTGAATTTTGTATAGCGAAGAGATTCCTAC 3462

QY 1835 TAGATTAATGATCTCAGGTGGAATCAGATGGCGCTTACACATGGTCAATTTTCTCTA 1894  
DB 3461 TAGATTAATGATCTCAGGTGGAATCAGATGGCGCTTACACATGGTCAATTTTCTCTA 3402

QY 1895 TTTTGCCGATTTATGATTCAGGTGGATACCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 1954  
DB 3401 TTTTGCCGATTTATGATTCAGGTGGATACCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 3342

QY 1955 TAGAAGCCCAACCGTCCAGATCGGATTTGTGCGCTTATCGTAGAAGAGCTGTTGCAAT 2014  
DB 3341 TAGAAGCCCAACCGTCCAGATCGGATTTGTGCGCTTATCGTAGAAGAGCTGTTGCAAT 3282

QY 2015 ATGACGACAAATGGCGTAATATTCGTGTGCAAGGACCGACAACTTCTACCACTCTGA 2074  
DB 3281 ATGACGACAAATGGCGTAATATTCGTGTGCAAGGACCGACAACTTCTACCACTCTGA 3222

QY 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTTGCTGTCATCGCATCTGTTATA 2134  
DB 3221 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATGTTGCTGTCATCGCATCTGTTATA 3162

QY 2135 CGACGAGCGTTCGGTCTTAACTGAACAGTTAAGCAATCAGATCTTCCCTTCAGGTTATG 2194

Db 3161 CGACGAGCGCTTCGGTCTTAACCTGAAGCAGTTAAGCAATCAGATCTTCCTTCAGGTTATG 3102  
Qy 2195 ACCATCTGTGCCAGTTCGTATGCTGCTCAACTTCGGACTCTGAGAACTTCTGGAA 2254  
Db 3101 ACCATCTGTGCCAGTTCGTATGCTGCTCAACTTCGGACTCTGAGAACTTCTGGAA 3042  
Qy 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACACCGATATATAGTGG 2314  
Db 3041 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACACCGATATATAGTGG 2982  
Qy 2315 ATGTGTCAAAACGCATACCATTTTGAA 2341  
Db 2981 ATGTGTCAAAACGCATACCATTTTGAA 2955  
RESULT 15  
US-10-385-415-4/c  
; Sequence 4, Application US/10385415  
; Publication No. US20040014158A1  
; GENERAL INFORMATION:  
; APPLICANT: Bacher, Adelbert  
; APPLICANT: Fischer, Maribus  
; TITLE OF INVENTION: PROTEIN CONJUGATES, METHOD, VECTORS, PROTEINS AND DNA FOR  
; TITLE OF INVENTION: PRODUCING THEM, THEIR USE AND MEDICAMENTS AND VACCINES CONTAININ  
; TITLE OF INVENTION: A CERTAIN QUANTITY OF SAID PROTEIN CONJUGATES  
; FILE REFERENCE: 9286.6CT  
; CURRENT APPLICATION NUMBER: US/10/385,415  
; CURRENT FILING DATE: 2003-03-10  
; PRIOR APPLICATION NUMBER: US 09/936,028  
; PRIOR FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: PCT/EP00/01899  
; PRIOR FILING DATE: 2000-03-03  
; PRIOR APPLICATION NUMBER: DE 19910102.7  
; PRIOR FILING DATE: 1999-03-08  
; NUMBER OF SEQ ID NOS: 154  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4  
; LENGTH: 5767  
; TYPE: DNA  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: p602-BS-LuSy Expression plasmid  
US-10-385-415-4  
Query Match 27.4%; Score 865.4; DB 7; Length 5767;  
Best Local Similarity 99.9%; Pred. No. 1.9e-206;  
Matches 866; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1475 TGCTTAGGAAGACGAGTTATTAAATAGCTGAATAAGAACGGTGTCTCTCCAAATATTCTTAT 1534  
Db 4286 TGCTTAGGAAGACGAGTTATTAAATAGCTGAATAAGAACGGTGTCTCTCCAAATATTCTTAT 4227  
Qy 1535 TTAGAAAAGCAAATCTAAAATTTATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA 1594  
Db 4226 TTAGAAAAGCAAATCTAAAATTTATCTGAAAAGGGAATGAGAATAGTGAATGGACCAATAA 4167  
Qy 1595 TAATGACTAGAGAAAGAAATGAGATTTGTCATGAATTAAGGAACGAATATTGGATA 1654  
Db 4166 TAATGACTAGAGAAAGAAATGAGATTTGTCATGAATTAAGGAACGAATATTGGATA 4107  
Qy 1655 AATATGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCTTTGGTCTCAGACTGATG 1714  
Db 4106 AATATGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCTTTGGTCTCAGACTGATG 4047  
Qy 1715 GGCCCTATTCCGATATTGAGATGATGTCATGTCTCAACAGAGAAAGCAGAGTTTCAGCC 1774  
Db 4046 GGCCCTATTCCGATATTGAGATGATGTCATGTCTCAACAGAGAAAGCAGAGTTTCAGCC 3987  
Qy 1775 ATGAATGGACAAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTATAGCGAAGATTCTAC 1834  
Db 3986 ATGAATGGACAAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTATAGCGAAGATTCTAC 3927  
Qy 1835 TAGATTATGCATCTCAGGTGGAATCAGATTGGCCGCTTTACACATGGTGGTCAATTTTCTCTA 1894

Db 3926 TAGATTATGCATCTCAGGTGGAATCAGATTGGCCGCTTTACACATGGTCAATTTTCTCTA 3867  
Qy 1895 TTTTGCCGATTTATGATTCAGGTGGATACCTTAGAGAAAGTGTATCAAACTGCTAAATCGG 1954  
Db 3866 TTTTGCCGATTTATGATTCAGGTGGATACCTTAGAGAAAGTGTATCCAACTGCTAAATCGG 3807  
Qy 1955 TAGAAGCCCAAAACGTTCCACGATGCGATTTGTGCCCTTATCGTAGAAGAGCTGTTTGAAT 2014  
Db 3806 TAGAAGCCCAAAACGTTCCACGATGCGATTTGTGCCCTTATCGTAGAAGAGCTGTTTGAAT 3747  
Qy 2015 ATGCAGGCAAAATGCGGTAATATTCTGTGTGCAAGGACCGCAACAACATTTCTACCATCCTTGA 2074  
Db 3746 ATGCAGGCAAAATGCGGTAATATTCTGTGTGCAAGGACCGCAACAACATTTCTACCATCCTTGA 3687  
Qy 2075 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTGTCTGCATCATCGCATCTGTTATA 2134  
Db 3686 CTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTGTCTGCATCATCGCATCTGTTATA 3627  
Qy 2135 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTTATG 2194  
Db 3626 CGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGATCTTCTTCAGGTTATG 3567  
Qy 2195 ACCATCTGTGCCAGTTTCGTAATGTCTGTGTCAACTTTCCGACTCTGAGAAAACCTTCTGGAAT 2254  
Db 3566 ACCATCTGTGCCAGTTTCGTAATGTCTGTGTCAACTTTCCGACTCTGAGAAAACCTTCTGGAAT 3507  
Qy 2255 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACAGGATATATAGTGG 2314  
Db 3506 CGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGACAGGATATATAGTGG 3447  
Qy 2315 ATGTGTCAAAACGCATACCATTTTGAA 2341  
Db 3446 ATGTGTCAAAACGCATACCATTTTGAA 3420

Search completed: February 12, 2006, 03:10:48  
Job time : 1633 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2006, 02:43:45 ; Search time 366 Seconds  
(without alignments)  
7753.446 Million cell updates/sec

Title: US-10-681-086-1  
Perfect score: 3156  
Sequence: 1 ggatccacgggttagcagc.....atcggtatgatgtcgaattc 3156

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 6240305 seqs, 449581930 residues

Total number of hits satisfying chosen parameters: 12480610

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications NA New:\*
- 1: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
  - 2: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
  - 3: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
  - 4: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
  - 5: /cgn2\_6/ptodata/1/pubpna/US05\_NEW\_PUB.seq:\*
  - 6: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
  - 7: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq:\*
  - 8: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
  - 9: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
  - 10: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq3:\*
  - 11: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq4:\*
  - 12: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*

Ptd. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description
c 1	762.4	24.2	4775	US-10-401-386B-62
c 2	286.6	9.1	3348	US-10-793-626-3541
c 3	286.6	9.1	3830	US-10-793-626-3934
c 4	286.6	9.1	3876	US-10-793-626-3895
5	285	9.0	999	US-10-793-626-539
6	285	9.0	999	US-10-793-626-2561
7	172	5.4	1470	US-11-000-688-675
8	114.6	3.6	3619	US-10-793-626-3335
9	113	3.6	1074	US-10-467-657-6969
10	107.8	3.4	1146	US-11-098-686-9321
11	107.8	3.4	1457619	US-11-098-686-8739
12	106.4	3.4	37507	US-10-522-037-2
13	100.6	3.2	3542	US-10-793-626-4459
14	95	3.0	1125	US-11-055-822-643
15	84	2.7	1970	US-11-136-527-2288
16	83	2.6	1134	US-10-793-626-2925
17	83	2.6	2955	US-10-793-626-3531
18	79.8	2.5	969	US-11-098-686-9320
19	79.8	2.5	1457619	US-11-098-686-8739
20	71.8	2.3	1049	US-11-136-527-174879
21	71.8	2.3	1049	US-11-136-527-174880

22	71.8	2.3	1049	11	US-11-136-527-174881	Sequence 174881,
23	71.8	2.3	1049	11	US-11-136-527-174905	Sequence 174905,
24	71.8	2.3	1049	11	US-11-136-527-174906	Sequence 174906,
25	71.8	2.3	1049	11	US-11-136-527-174907	Sequence 174907,
26	71.8	2.3	1049	11	US-11-136-527-174908	Sequence 174908,
27	71.8	2.3	1049	11	US-11-136-527-174909	Sequence 174909,
28	71.8	2.3	1049	11	US-11-136-527-174910	Sequence 174910,
29	71.8	2.3	1049	11	US-11-136-527-174911	Sequence 174911,
30	71.8	2.3	1049	11	US-11-136-527-174912	Sequence 174912,
31	71.8	2.3	1049	11	US-11-136-527-174913	Sequence 174913,
32	71.8	2.3	1049	11	US-11-136-527-174914	Sequence 174914,
33	71.8	2.3	1049	11	US-11-136-527-174915	Sequence 174915,
34	71.8	2.3	1049	11	US-11-136-527-174916	Sequence 174916,
35	71.8	2.3	1049	11	US-11-136-527-174917	Sequence 174917,
36	71.8	2.3	1049	11	US-11-136-527-174918	Sequence 174918,
37	71.8	2.3	1049	11	US-11-136-527-174919	Sequence 174919,
38	71.8	2.3	1049	11	US-11-128-061-3592	Sequence 3592, Ap
39	71.8	2.3	1049	11	US-11-128-061-3593	Sequence 3593, Ap
40	71.8	2.3	1049	11	US-11-128-061-3594	Sequence 3594, Ap
41	71.8	2.3	1049	11	US-11-128-049-3592	Sequence 3592, Ap
42	71.8	2.3	1049	11	US-11-128-049-3593	Sequence 3593, Ap
43	71.8	2.3	1049	11	US-11-128-049-3594	Sequence 3594, Ap
44	63.2	2.0	333	11	US-11-128-061-7236	Sequence 7236, Ap
45	63.2	2.0	333	11	US-11-128-049-7236	Sequence 7236, Ap

ALIGNMENTS

RESULT 1  
US-10-401-386B-62/c  
; Sequence 62, Application US/10401386B  
; Publication No. US20050261213A1  
; GENERAL INFORMATION:  
; APPLICANT: Patrick Branigan  
; APPLICANT: Theresa J Goletz  
; APPLICANT: David M Knight  
; APPLICANT: Stephen G McCarthy  
; APPLICANT: Bernard J Scallion  
; APPLICANT: Linda A Snyder  
; TITLE OF INVENTION: Nucleic Acid Compositions and Methods  
; FILE REFERENCE: CEN 310CIP  
; CURRENT APPLICATION NUMBER: US/10/401.386B  
; CURRENT FILING DATE: 2003-03-28  
; PRIOR APPLICATION NUMBER: 10/247,203  
; PRIOR FILING DATE: 2002-09-19  
; PRIOR APPLICATION NUMBER: 60/328,371  
; PRIOR FILING DATE: 2001-10-10  
; NUMBER OF SEQ ID NOS: 81  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 62  
; LENGTH: 4775  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)...(1177)  
; OTHER INFORMATION: Kanamycin resistance  
; FEATURE:  
; NAME/KEY: rep origin  
; LOCATION: (1178)...(2129)  
; FEATURE:  
; NAME/KEY: enhancer  
; LOCATION: (2371)...(2663)  
; OTHER INFORMATION: Rous Sarcoma Virus enhancer  
; FEATURE:  
; NAME/KEY: promoter  
; LOCATION: (2766)...(3508)  
; OTHER INFORMATION: HCMV  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (3509)...(4533)

```
; OTHER INFORMATION: PSA CDNA
; FEATURE:
; NAME/KEY: polyA site
; LOCATION: (4580)...(4775)
; OTHER INFORMATION: SV40
US-10-401-386B-62

Query Match      24.2%; Score 762.4; DB 7; Length 4775;
Best Local Similarity 96.8%; Pred. No. 7.8e-182;
Matches 778; Conservative 0; Mismatches 26; Indels 0; Gaps 0;

Qy 1564 AAGGGAATGAGATAGTGAATGGAACCAATAATAATGACTAGAGAAAGAAATGAAGATT 1623
Db      |||
Qy 1074 AATCAAGGGGTGTTATGAACGGCAATAATAATGACTAGAGAAAGAAATGAAGATT 1015
Db      |||
Qy 1624 GTTCATGAATTAAGGAACGAATATTGGATAAATATGGGGATGATGTTAAGGCTATTGGT 1683
Db      |||
Qy 1014 GTTCATGAATTAAGGAACGAATATTGGATAAATATGGGGATGATGTTAAGGCTATTGGT 955
Db      |||
Qy 1684 GTTTATGGCTCTCTTGGTCGTGAGACTGATCGGCCCTATTTCGGATATTGAGATGATGTT 1743
Db      |||
Qy 954 GTTTATGGCTCTCTTGGTCGTGAGACTGATCGGCCCTATTTCGGATATTGAGATGATGTT 895
Db      |||
Qy 1744 GTCATGTCAACAGAGGAACGAGAGTTTCAGCCATGAATGACACCCGGTGAAGAGGTG 1803
Db      |||
Qy 894 GTCATGTCAACAGAGGAACGAGAGTTTCAGCCATGAATGACACCCGGTGAAGAGGTG 835
Db      |||
Qy 1804 GAAGTGAATTTTGATAGGCAAGAGATTCTACTAGATTATGATCTCAGGTGGAATFCAGAT 1863
Db      |||
Qy 834 GAAGTGAATTTTGATAGGCAAGAGATTCTACTAGATTATGATCTCAGGTGGAATFCAGAT 775
Db      |||
Qy 1864 TGGCGCTTACACATGGTCAATTTTCTCTATTTTCCGATTTATGATTCAGGTGGATAC 1923
Db      |||
Qy 774 TGGCGCTTACACATGGTCAATTTTCTCTATTTTCCGATTTATGATTCAGGTGGATAC 715
Db      |||
Qy 1924 TTAGAGAAGTGTATCAAACTGCTAAATCGGTAGAACGCCAAAGCTTCCAGATGCCGATT 1983
Db      |||
Qy 714 TTAGAGAAGTGTATCAAACTGCTAAATCGGTAGAACGCCAAAGCTTCCAGATGCCGATT 655
Db      |||
Qy 1984 TGTGCCCTTATCGTAGAAGAGCTGTTTGAATATGACGCAAAATGCGTGAATATTTCGTGTG 2043
Db      |||
Qy 654 TGTGCCCTTATCGTAGAAGAGCTGTTTGAATATGACGCAAAATGCGTGAATATTTCGTGTG 595
Db      |||
Qy 2044 CAAGACCGCAACAACTTTTACCCTCTGACTGTACAGTACAGTACAGTACAGTACAGTACAGT 2103
Db      |||
Qy 594 CAAGACCGCAACAACTTTTACCCTCTGACTGTACAGTACAGTACAGTACAGTACAGTACAGT 535
Db      |||
Qy 2104 TTGATTGGTCTGCAATCGCATCTGTTTATACGACGAGCGCTTCGGTCTTAACTGAAGCA 2163
Db      |||
Qy 534 TTGATTGGTCTGCAATCGCATCTGTTTATACGACGAGCGCTTCGGTCTTAACTGAAGCA 475
Db      |||
Qy 2164 GTTAAGCAATCAGATCTTCCCTTACGTTATGACCAATCTGTGCCAGTTCGTAATGTCGTGT 2223
Db      |||
Qy 474 GTTAAGCAATCAGATCTTCCCTTACGTTATGACCAATCTGTGCCAGTTCGTAATGTCGTGT 415
Db      |||
Qy 2224 CAATTTCCGACTCTGAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAG 2283
Db      |||
Qy 414 CAATTTCCGACTCTGAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAG 355
Db      |||
Qy 2284 GAGTGGACAGAACGACACGGATATATAGTGGATGTGTCAAAACGCAATACCAATTTTGAATT 2343
Db      |||
Qy 354 GAGTGGACAGAACGACACGGATATATAGTGGATGTGTCAAAACGCAATACCAATTTTGAACG 295
Db      |||
Qy 2344 CGAAAGCCCGAATGAGTCTTACC 2367
Db      |||
Qy 294 CAGAAATTTGGTTAATTGGTTGTAAC 271
Db      |||
```

## RESULT 2

```
US-10-793-626-3541/c
; Sequence 3541, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
```

```
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; PRIOR FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3541
; LENGTH: 3348
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3541
```

```
Query Match      9.1%; Score 286.6; DB 7; Length 3348;
Best Local Similarity 61.8%; Pred. No. 4.2e-62;
Matches 473; Conservative 0; Mismatches 289; Indels 3; Gaps 1;

Qy 2371 TGGTGAATTAAGGAACCCCTGCTTGAAGCGCGAAGCGGCGCACGATCTGAATATCGGCA 2430
Db      |||
Qy 2492 TTGTTGAACAGGACCAAAATTAAGAAGCGCTCAAGTGGCAACTGAAAAATCAAAATCGGTA 2433
Db      |||
Qy 2431 CATATTGTCGTGGCAAGCGGAGAGTCCGCTTAACAGAGAGTGGATCAGTCCGTAG 2490
Db      |||
Qy 2432 CATACGTATTTGTTATGAGTGGTAGAGTCTTAGTAAACAGAGAGTGCATCATATTGCG 2373
Db      |||
Qy 2491 ATGCGGTTACAGGAATTAAGAAGACGTA---TGACCTGAAGATTTGTGCATGCTTGGAC 2547
Db      |||
Qy 2372 AAACAGTAGAAGATATTAAAGAGATACACCCACNACTAAAGATTTGTGCGTCTAGGAT 2313
Db      |||
Qy 2548 TGTGAAGCCAGAGCAGCGGCTCAAAGATGAGAGATAGACCGCTATATATCATATA 2607
Db      |||
Qy 2312 TAACGAAGAAGAACACAGGCTAAAAAATTAAGGCTGCTGGTGCATCGTTATAATCATATA 2253
Db      |||
Qy 2608 ATTTGAATAGCTACAGAGAAACCATTCAAACATCAACCTCACATACATACATGACA 2667
Db      |||
Qy 2252 ATTTAAATACGAGTGAGCGTTATCAAGATGAAGTAGTAACCTACACATACATATGAGGATA 2193
Db      |||
Qy 2668 GAGTCAATACGTTGAATTCGAAATCGCAAAAGAAATCGGGGCTGTCTCCGTGTTTCAGCGCCCATTA 2727
Db      |||
Qy 2192 GAGTGAATACGTTGAATTCGAAATGATGAAGATTAATAATTTCTCTTGTTCAGTGTGATAT 2133
Db      |||
Qy 2728 TCGGGATGAAGAGAGACGAACACAGATGTCATGACATCGCCAAAGCTTGAAGGCTCTTG 2787
Db      |||
Qy 2132 GTGGTATGGGAGAGTCCGAATGAGGACATTTATGATATGGCATTTTGCTTTAAGAGCCATCG 2073
Db      |||
Qy 2788 ACGCGGATTCATTCCTGTGAAATTTTTCGATGCAATTTGATGCAACCGCTTGAAGCGG 2847
Db      |||
Qy 2072 ATGCTGATAGCATTCCTTATTAATTTTTCATCATCTTAAAGGAACTAAATTTTCGTGGAT 2013
Db      |||
Qy 2848 TCACGAATTTAAACCCCTGTTATTTTAAAGTGCTGGGCTGTTCGCTTTTATCAATC 2907
Db      |||
Qy 2012 TAGATTTATTTGTCACCAATGAATTTTAAAGATTAATAGCGATGTTTAGGTTATCAATC 1953
Db      |||
Qy 2908 CATCAAAAGAAATTCGCATTTTCGGAGAGAGAGAGTCAATCTCCGACATGTCAGGCAAT 2967
Db      |||
Qy 1952 CAACAAAGAAATTCGAATTCGAGTGGAGGAGGTAATCTACGTTTCAATTAACAACCAAC 1893
Db      |||
Qy 2968 TAGGGCTTTACGGCGCAAACTCCATTTTTCGAGGACTACTTAAACACTGCGCGGCAAG 3027
Db      |||
Qy 1892 TCGCATTTGAAGCGGCTAATTTCAATTTTTCAGGAGATTACTTAAATTACAGGCGGTCAAC 1833
Db      |||
Qy 3028 AGGAGACGAGAGATCATATAAATCTGAGTGAATTTAGGCTTTGAGATTCAGTCCGAAG 3087
Db      |||
Qy 1832 CGATGAGGAGATTAATCGCATGATGAAGATTTAGGCTTTGAAATCGACAGTTAAATTA 1773
Db      |||
Qy 3088 AAATGAAGGCTAGTTTAAGTGGCAAAAGCTGAAAGAAATCAATAAA 3132
Db      |||
Qy 1772 TAAATTAATAATATTTTCAATAAATGTAATAATATAGTTTAAATAAA 1728
Db      |||
```

## RESULT 3

US-10-793-626-3934/c  
; Sequence 3934, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3934  
; LENGTH: 3830  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-3934

	Query Match	9.1%;	Score 286.6;	DB 7;	Length 3830;
	Best Local Similarity	61.8%;	Pred. No. 4.4e-62;		
	Matches 473;	Conservative 0;	Mismatches 289;	Indels 3;	Gaps 1;
QY	2371	TGTTGAATAAGGAAACGCTCTTGAAGCGCGGAGCGGCGACGATCTGAATATCGGCA	2430		
DB	887	TTGTTGAACAGGACCAAAATTAAGAAGCGCTCAAGTGGCAACTGAAATCAAATCGGTA	828		
QY	2431	CATATTGTATCGTGGCAAGCGGCGAGGTCGGTCTAACAGAGAGTGATCAGGTCGTAG	2490		
DB	827	CATCTGTATTGTATGAGTGGTAGAGTCTTAGTAAACAGAGTGCATCATATTTCGG	768		
QY	2491	ATCGCGTTTCAGGAAATTAAGAGACGTA---TGGACTGAAAGATTGTGCATGCTCTGGAC	2547		
DB	767	AAACAGTAGAGATATTAAAGAGATACACCCACAACATAAGATTGTGCGTCTAGGAT	708		
QY	2548	TGTTGAAGCCAGAGCGGCGGAGCGGCTCAAAGATGAGAGTAGACCGCTATAATCAT	2607		
DB	707	TAACGAAAGAGAACAGCGCTAAAGAAATTAAGGCTGCTGTCGATCGTTATAATCAT	648		
QY	2608	ATTGTAATACGTACAGAGAAACCATTCAAACATCAACCTCACATACATACATGATGACA	2667		
DB	647	ATTAAATACGTGAGCGTTATCAGATGAAGTAGTAACATACATACATATGAGGATA	588		
QY	2668	GAGTCAATACGTTGAAATCGCAAAAGAAATCGGGGCTGTCTCGGTGTTCAAGCGCCATTA	2727		
DB	587	GAGTGAATACGTTGAAATGATGAAGATAAATAATTTCTCTCTGTTCAAGTGTGAT	528		
QY	2728	TCGGGATGAAGAGAGAGCAAAACAGATGTCATTCGATCGCAAAAGCTTGAAGCGCTTTG	2787		
DB	527	GTGTTATGGAGAGTCAATGAGGACATTTATGATATGCGCATTTGCTTTAAGAGCCATCG	468		
QY	2788	ACGGGATTCATCTCTGTGAATTTTTCATGCAATGATGATGCGCGCTTGAAGCGG	2847		
DB	467	ATGCTGATAGCATTCCTATTAAATTTTATCATCTCTATTAAAGGAACCTAAATTTGGTGGAT	408		
QY	2848	TCAACGAATTAACCCCGCTGATTGTTTAAAGTGTGGCGCTGTTCCCGTTTATCAATC	2907		
DB	407	TAGATTATTGTCACCAATGAATGTTTAAAGATTAAGCATGTTTAGGTTAATCAATC	348		
QY	2908	CATCAAAAGAAATTCGCAATTCGGAGGAGAGAGGTCAATCTCCGACATTCAGCCAT	2967		
DB	347	CAACAAAGAAATTCGAAATTCGAGGTGAGCGGAGGTAAATCTACGTTTCATTACAAACCAC	288		
QY	2968	TAGGGCTTTAGCGGCGCAACTCCATTTTGTTCGAGACTACTTAAACACTGCGCGGCAAG	3027		
DB	287	TCGCATTGAAGCGGCTAATTCATTTTGTAGGAGATTACTTAATTAACAGCGGCTCAAC	228		

QY	3028	AGAGAGCGGAGATCAATAAAATGCTGAGTGATTTAGGCTTTGAAGTTGAATCAAGTCGAAG	3087		
DB	227	CGAATGAGGAAGATTATCGCATGATTGAAGATTTAGGTTTGAATCGCAGTTAAATTA	168		
QY	3088	AAATGAGGCTAGTTTAAAGTGCAGAAAGCTGAAGATCAATAAA	3132		
DB	167	TAAATTAATAATATTTTCAATAAATGTAATAATATAGTTTAAATAAA	123		

## RESULT 4

US-10-793-626-3895/c  
; Sequence 3895, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3895  
; LENGTH: 3876  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-3895

	Query Match	9.1%;	Score 286.6;	DB 7;	Length 3876;
	Best Local Similarity	61.8%;	Pred. No. 4.4e-62;		
	Matches 473;	Conservative 0;	Mismatches 289;	Indels 3;	Gaps 1;
QY	2371	TGTTGAATAAGGAAACGCTCTTGAAGCGGCGGAGCGGCGACGATCTGAATATCGGCA	2430		
DB	3117	TTGTTGAACAGGACCAAAATTAAGAAGCGCTCAAGTGGCAACTGAAATCAAATCGGTA	3058		
QY	2431	CATATTGTATCGTGGCAAGCGGCGAGGTCGGTCTAACAGAGAGTGATCAGGTCGTAG	2490		
DB	3057	CATCTGTATTGTATGAGTGGTAGAGTCTTAGTAAACAGAGATGCGATCATATTTCGG	2998		
QY	2491	ATCGCGTTTCAGGAAATTAAGAGACGTA---TGGACTGAAAGATTGTGCATGCTCTGGAC	2547		
DB	2997	AAACAGTAGAGATATTAAAGAGATACACCCACAACATAAGATTGTGCGTCTAGGAT	2938		
QY	2548	TGTTGAAGCCAGAGCGGCGGAGCGGCTCAAAGATGAGAGTAGACCGCTATAATCAT	2607		
DB	2937	TAACGAAAGAGAACAGCGCTAAAGAAATTAAGGCTGCTGTCGATCGTTATAATCAT	2878		
QY	2608	ATTGTAATACGTACAGAGAAACCATTCAAACATCAACCTCACATACATACATGATGACA	2667		
DB	2877	ATTAAATACGTGAGCGTTATCAGATGAAGTAGTAACATACATACATATGAGGATA	2818		
QY	2668	GAGTCAATACGTTGAAATCGCAAAAGAAATCGGGGCTGTCTCGGTGTTCAAGCGCCATTA	2727		
DB	2817	GAGTGAATACGTTGAAATGATGAAGATAAATAATTTCTCTCTGTTCAAGTGTGAT	2758		
QY	2728	TCGGGATGAAGAGAGAGCAAAACAGGATGTCATTCGATCGCAAAAGCTTGAAGCGCTTTG	2787		
DB	2757	GTGTTATGGAGAGTCAATGAGGACATTTATGATATGCGCATTTGCTTTAAGAGCCATCG	2698		
QY	2788	ACGGGATTCATCTCTGTGAATTTTTCATGCAATGATGATGCGCGCTTGAAGCGG	2847		
DB	2697	ATGCTGATAGCATTCCTATTAAATTTTATCATCTCTATTAAAGGAACCTAAATTTGGTGGAT	2638		
QY	2848	TCAACGAATTAACCCCGCTGATTGTTTAAAGTGTGGCGCTGTTCCCGTTTATCAATC	2907		
DB	2637	TAGATTATTGTCACCAATGAATGTTTAAAGATTAAGCATGTTTAGGTTAATCAATC	2578		
QY	2908	CATCAAAAGAAATTCGCAATTCGGAGGAGAGAGGTCAATCTCCGACATTCAGCCAT	2967		

Db 2577 CAA CAA A A A A A A A T T C G A A T T C A G G T G C A C G G G A G G T A A A T C T A C G T T C A T T A C A A C C A C 2518  
Qy 2568 T A G G C G T T A C G C G C A A A C T C C A T T T T G T C G A G A C T A C T T A A C A A C T C C C G G G C A A G 3027  
Db 2517 T C G C A T T G A A A G C G G C T A A T T C A A T T T T G T A G A G A T T A C T T A A T T A C A G C G G T C A A C 2458  
Qy 3028 A G G A G C G G A G A T C A T A A A A T G C T G A G T G A T T T A G G C T T T G A A G T T G A A T C A G T C G A A G 3087  
Db 2457 C G A A T G A G A G A T T A T C G C A T G A T T G A G A T T T A G G T T T G A A T C G A C A G I T T A A A T T A 2398  
Qy 3088 A A A T G A A G C T A G T T T A A G T C G A A A A G C T G A A A G A A T C A A T A A A 3132  
Db 2397 T A A T T A A A A T A T T T T C A A T A A A T G T A A A T A T A G T T T T A A A T A A A 2353

## RESULT 5

US-10-793-626-539  
; Sequence 539, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; PRIOR FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 539  
; LENGTH: 999  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-539

Query Match 9.0%; Score 285; DB 7; Length 999;  
Best Local Similarity 63.6%; Pred. No. 6.5e-62;  
Matches 451; Conservative 0; Mismatches 255; Indels 3; Gaps 1;  
Qy 2371 T G T G A A T A A G G A A A C G T C T T C A A G C G C G A A G C G G C A C G A T C T G A A T A T C G C A 2430  
Db 284 T T G T T G A A C A G A C C A A A T T A A A G A G C G C T C A A G T G G C A A A T C A A A T C G G T A 343  
Qy 2431 C A T A T T G T A T C G T G C A A G C G C A G A G T C C G T C T A A C A G A A A G T G A T C A G T C G T A G 2490  
Db 344 C A T A C T G T A T T G T T A G T G T A G A G T C C T A G T A A C A G A A G T C G A T C A T A T T T G C G 403  
Qy 2491 A T G C G G T T C A G A A A T T A A G A G A C G T A --- T G A C T G A A G A T T T G T C A T G T C T T G G A C 2547  
Db 404 A A A C A G T A G A A G A T T A A A A A A A T A C A C C A C A A C T A A A G A T T T G T C G T C T T A G A T 463  
Qy 2548 T G T T G A A C C A G A C A G C G A A G C G G C T C A A A G A T G C A G A G T A G A C C G T A T A A T C A T A 2607  
Db 464 T A A C G A A A G A A C A G C T A A A A A T T A A A G G C T G C T G T G T C G A T C G T T A T A A T C A T A 523  
Qy 2608 A T T T G A A T A C G T C A C A G A A A C C A T T C A A A C A T C A C A C C T C A C A T A C A T A C A T A G C A 2667  
Db 524 A T T T A A A T A C A G T G A C G T T A T C A C A T G A A G T A G T A A C T A C A C A T A C A T A T A G A G A T A 583  
Qy 2668 G A G T C A A T A C G G T T G A A A T C G A A A A G A A T C G G G C T G T C T C C G T G T T C A G C C C A T T A 2727  
Db 584 G A G T G A A T A C G G T T G A A A T G A T A A A G A T A A T A T T T C T C T T G T T C A G G T G T G A T A T 643  
Qy 2728 T C G G G A T G A A G G A C G A A A C A G A T G C A T T G A C T C G C C A A A A G C T T G A A G C T C T T G 2787  
Db 644 G T G G T A T G G G A G A T C G A A T G A G G A C A T T A T T G A T A T G G C A T T T G C T T T A A G A C C A T C G 703  
Qy 2788 A C G C G G A T T C C A T T C C T G T G A A T T T T T G C A T G C A A T T G A T T G G C A C C G C G T T A G A A G C G 2847

Db 704 A T G C T A T A G C A T T C C T A T T A A T T T T T T A C A T C C T A T T A A A G A A C T A A A A T T T G T G G A T 763  
Qy 2848 T C A A C G A A T T A A A C C G C T C T A T T T G T T T A A A G T G C G C G C T C T C C G T T T T A T C A A T C 2907  
Db 764 T A G A T T A T T G T C A C C A A T G A A A T G T T T A A A A T T A T A G C A T T T A G G T T A T C A A T C 823  
Qy 2908 C A T C A A A A G A A A T T C G C A T T T C C G A G A A G A G A G G T C A A T C T C C G A C A C A T T G C A G C C A T 2967  
Db 824 C A A C A A A A G A A A T T C G A A T T G C A G G T G C A C G G A G G T A A A T C T A C G T T C A T T A C A C C A C 883  
Qy 2968 T A G G C G T T A C G C G C A A A C T C C A T T T T G T C G A G A C T A C T T A A C A A C T G C C C G G C A A G 3027  
Db 884 T C G C A T T G A A A G C G C T A A T T C A A T T T T T G T A G A G A T T A C T T A A T T A C A G C G G T C A A C 943  
Qy 3028 A G G A G C G G A G A T C A T A A A A T G C T G A T G A T T T A G G C T T T G A A G T T G A 3076  
Db 944 C G A A T G A G A A G A T T A T C G C A T G A T T A G G G T T T G A A A T C G A 992

## RESULT 6

US-10-793-626-2561  
; Sequence 2561, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2561  
; LENGTH: 999  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-2561

Query Match 9.0%; Score 285; DB 7; Length 999;  
Best Local Similarity 63.6%; Pred. No. 6.5e-62;  
Matches 451; Conservative 0; Mismatches 255; Indels 3; Gaps 1;  
Qy 2371 T G T G A A T A A G G A A A C G C T C T T G A A G C G C G A A G C G G C A C G A T C T G A A T A T C G C A 2430  
Db 284 T T G T T G A A C A G A C C A A A T T A A A G A G C G C T C A A G T G G C A A A T C A A A T C G G T A 343  
Qy 2431 C A T A T T G T A T C G T G C A A G C G C A G A G T C C G T C T A A C A G A A A G T G A T C A G T C G T A G 2490  
Db 344 C A T A C T G T A T T G T T A G T G T A G A G T C C T A G T A A C A G A A G T C G A T C A T A T T T G C G 403  
Qy 2491 A T G C G G T T C A G A A A T T A A A G A G A C G T A --- T G A C T G A A G A T T T G T C A T G T C T T G G A C 2547  
Db 404 A A A C A G T A G A A G A T T A A A A A G A T A C A C C A C A A C T A A A G A T T T G T G G T G C T T A G A T 463  
Qy 2548 T G T T G A A C C A G A C A G C G G C T C A A A G A T G C A G A G T A G A C C G C T A T A A T C A T A 2607  
Db 464 T A A C G A A A G A A C A G C G T A A A A A A T T A A A G G C T G C T G T G T C G A T C G T T A T A A T C A T A 523  
Qy 2608 A T T T G A A T A C G T C A C A G A A A C C A T T C A A A C A C T C A C A C C T C A C A T A C A T A C A T A G C A 2667  
Db 524 A T T T A A A T A C A G T G A C G T T A T C A C A T G A A G T A G T A A C T A C A C A T A C A T A T A G A G A T A 583  
Qy 2668 G A G T C A A T A C G G T T G A A A T C G A A A A G A A T C G G G C T G T C T C C G T G T T C A G C C C A T T A 2727  
Db 584 G A G T G A A T A C G G T T G A A A T G A T A A A G A T A A T A T T T C T C T T G T T C A G G T G T G A T A T 643  
Qy 2728 T C G G G A T G A A G G A C A A A C A G A T G C A T T G A C A T C G C C A A A A G C T T G A A G C T C T T G 2787  
Db 644 G T G G T A T G G G A G A T C G A A T G A G G A C A T T A T T G A T A T G G C A T T T G C T T T A A G A C C A T C G 703

QY	2788	ACGGGATTCATTCCTGTGAATTTTTTGCATCCAATTGATGCACGCCGTTAGAAGCG	2847
Db	704	ATGCTGATAGCATTCCTATTAAATTTTTTACATCCTATTAAAGGAATTAATTTGGTGGAT	763
QY	2848	TCAACGAATTAACACCGCTGTATGTGTTTAAAAGTGCTGGCGCTGTTCCGTTTTATCAATC	2907
Db	764	TAGATTATTGTGTCACCAATGAAATGTTTAAGAAATTATAGCATGTTTAGGTAAATCAATC	823
QY	2908	CATCAAAAGAAATTCGATTTCCGGAGGAGAGAGGTCATCTCCGCACATTCGACGCCAT	2967
Db	824	CAACAAAGAAATTCGAAATTCGAGTGGACGGGAGTAAATCTACGTTCAATCAACCCAC	883
QY	2968	TAGGGCTTTACGGCGMAATCATATTTTTGTGCGAGACTACTTAAACATGCCGGCGGAAG	3027
Db	884	TCGCATTGAAGCGGCTAATTCAAATTTTTGTAGAGAACTACTTAATTACAGCGGGTCAAC	943
QY	3028	AGGAGACGGAGGATCATAAATCGTAGTGATTTAGGCTTTGAAGTTGA	3076
Db	944	CGAATCAGGAGAAATTAATCGCATGATTAAGAAATTTAGGTTTGAATTCGA	992

## RESULT 7

```

US-11-000-688-675
; Sequence 675, Application US/11000688
; Publication No. US20050287544A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, Francois
; APPLICANT: HOULGATTE, Remi
; APPLICANT: BIRNBAUM, Daniel
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF COLON CANCER WITH DNA ARRAYS
; FILE REFERENCE: 1423-R-03
; CURRENT APPLICATION NUMBER: US/11/000,688
; CURRENT FILING DATE: 2004-12-01
; PRIOR APPLICATION NUMBER: US 60/525,987
; PRIOR FILING DATE: 2003-12-01
; NUMBER OF SEQ ID NOS: 1596
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 675
; LENGTH: 1470
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial sequences:primer
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1470)
; OTHER INFORMATION: glycine c-acetyltransferase
; OTHER INFORMATION: (2-amino-3-ketobutyrate coenzyme a ligase) (GCAT)
; OTHER INFORMATION: gene.
US-11-000-688-675

```

Query Match	5.4%	Score 172;	DB 11;	Length 1470;
Best Local Similarity	49.7%;	Pred. No. 2.1e-33;		
Matches 473;	Conservative 0;	Mismatches 470;	Indels 9;	Gaps 1;
Qy	158	CAAAACAATTATTTAGGGCTCGCAAGCCATAGACGTTTTGATCGATGACAGCCCAACACGCAAT	217	
Db	235	CCAACACTACCTGGGCTCGAGCAGCCACCTCGAGGTGATCCAGGCAGGTCTGCAGGCTC	294	
Qy	218	TGCGCAATTTGGGACAGGAAGCAGCGGTTCACGTTTTAACACAGGCCAATTCGGTCTGGC	277	
Db	295	TGGAGGAGTTTGAGAGCTGGGCTCAGCTCGGTCGGCTTTATCTGTGGAAACCCAGAGCATCC	354	
Qy	278	ATGAAAAGCTAGAAAAAGAGATTGCCAGCTTTTAAACTGCACAGAAGCGGCCCTGCTGTTTT	337	
Db	355	ACAAGATCTAGAAAGCAAAAATAGSCCGCTTCCACACGCGGAGGATGCCATCCTCTATC	414	
Qy	338	CGAGCGGTTACTTGGCCCAATGTCGGTCTCCTTTTCATCCTTCCACAGAAAAGGAGATGTCA	397	
Db	415	CCAGCTGTTATAGCGCCAAACGCGGGCTCTTTGAGGGCCCTTGCTGACCCCAAGAGACCGAG	474	
Qy	398	TTTTAAGTCACCAAGCTCAATCATGCAAGATATGATCGACGGCTCGCGACTTTCTAAGGCTG	457	

```

RESULT 8
US-10-793-626-3335
; Sequence 3335, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: F03480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3335
; LENGTH: 3619
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3335

Query Match          3.6%; Score 114.6; DB 7; Length 3619;
Best Local Similarity 62.6%; Pred. No. 8,7e-19;

```

Query Match 3.6%; Score 114.6; DB 7; Length 3619;  
Best Local Similarity 62.6%; Pred. No. 8.7e-19;







GENERAL INFORMATION:  
; APPLICANT: LIBRAGEN  
; TITLE OF INVENTION: Method for the expression of unknown environmental DNA into adapted cells.  
; TITLE OF INVENTION: cells.  
; FILE REFERENCE: B0149WO  
; CURRENT APPLICATION NUMBER: US/10/522,037  
; CURRENT FILING DATE: 2005-01-24  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2 37507  
; LENGTH: 37507  
; TYPE: DNA  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: DNA sequence of clone FS3-135.  
US-10-522-037-2

Query Match 3.4%; Score 106.4; DB 7; Length 37507;  
Best Local Similarity 46.7%; Pred. No. 2.6e-16;  
Matches 452; Conservative 0; Mismatches 501; Indels 15; Gaps 3;  
  
Qy 150 CTGGTCTCAACAACTATTATAGGCTCGCAAGCATAGAGCTTTGATCGATGAGCCCA 209  
Db 1109 CTTCTGCGCCAAATACTACCTCGGCTGGCCAAACACCCCGCTGTCGGCGCGGAA 1168  
  
Qy 210 AACAGCATTCAGCAATTTGGACAGGAAGCAGCGGTTTACGTTTAAACAGCAGCAATTC 269  
Db 1169 GGAGAGCTTCAGCGCGAGGCTACGGCATGGCGTGGCTTCTCATCTCGGGAGCGCA 1228  
  
Qy 270 GGTCTGGCATGAAAGCTAGAAAAGAAATTTGCCAGCTTTAAATCGACAGAGCGGCCCT 329  
Db 1229 CGCGTGCAACCGAGCTGGAGACGCGCTCGCGCTTCTCGGCACCGAGGACGGAT 1288  
  
Qy 330 GCTGTTTCGAGCGGTTACTTTGGCAATGTGGTGTCTTTTCATCTTCCAGAAAAGGA 389  
Db 1289 CCTGTACGGCTCTGCTTCGACGCGCAACGGCGGCTCTTCGAGACGCTGTGGCGGAGGA 1348  
  
Qy 390 AGATGTCTATTTAAGTGACAGCTCAATCATGCAAGTATGATCGACGCTGCCACTTTC 449  
Db 1349 GGACGGGTGATTCGGACGCGCTGACACCGCTTCGATCATCGACGCGTGGCGCTTC 1408  
  
Qy 450 TAAGCTGATACAGTGTGTTATCGGCATPATTTGATATGATGATCTTGAAACAAAGCTGAA 509  
Db 1409 GAAGCGAAGCGCTTCGCTACGACAAACACGACATGGCGAGCTTCGAGGCCAGCTGAA 1468  
  
Qy 510 TGAACACAGCGTTATCAGCGCGTT-----TTATGCTACACAGCGAGTTTACGAT 563  
Db 1469 GCGCGCGACCGCGCGCGCGCTACAGATGATCGCCACCGACGCGGTGTTCTCGAT 1528  
  
Qy 564 GGATGGCAACAATCGCCCTCTTTGATCAGATCATCTCACTTCGGAACGCTATCATGCTT 623  
Db 1529 GGACGGATCGTCGGGAACCTTGGGGCCATCTCGAGCTGCGGGCGCTACACAGCGAT 1588  
  
Qy 624 CGTGGTCTGATGATGCCCAACGACAGGAGTTTGGCGGATTCGGGACAAAGAACGAG 683  
Db 1589 GGTGATGTGACGACTCGACGCGCTGCTTATGGCGCGCACGCGGGGGAACGCC 1648  
  
Qy 684 TGAATACTTTGGTGT-----TTGTCCGACATTTGTTATCGGACCTTTAAGCAAGC--- 734  
Db 1649 CGAGCATTTGGGTGTTCGAAGGGAAGGTGGACATCTCACCGGCAACGCTCGGCAAGGCCCT 1708  
  
Qy 735 TGTTCGGCGGAAAGGAGGTTTGGCGGAGGATCAGCGTCTTCATCGACTTTTTCGTTGAA 794  
Db 1709 GGGCGGCGCTTCGGCGGCTACAGCGGGGCAAGCGGAGGTGGCTGGCTTCGCA 1768  
  
Qy 795 CCATGCCAGAAATTTATCTTTTCAAAACCGCTATTCCCGCCAGCGAGTGTGGCGTGTCA 854  
Db 1769 CCGCTCGCGCCCTATCTCTCTCCACACGCTGATGCCCGCATCGCGGCGCGTCTGCT 1828  
  
Qy 855 CGAGCTTTCAACATCATTTGAAGCCAGCAGGAAACAGCAGCTTTTATTTCTTATAT 914  
Db 1829 CAAGGTGCTCGATCTCTCTCGAGGGCGGCGGAGCTGCGCGGAGCTCGCGCAACGC 1888

Qy 915 CAGCATGATCAGAACCAAGCTCTGAAGAATATATGGTTATGTGTGAAGAGGATCACACAC 974  
Db 1889 CCGCCACTTCGCGCGGAGATGACGCGCTTCAGCTGCGCGCGCGCGGACCATCC 1948  
  
Qy 975 GATTATTCCTGTAGTCAATTGGCGATGCCCAATAAACCGTCTTATTTGCTGAAAACCTGCA 1034  
Db 1949 GATCATCCGCTGATGCTGGCGAGGCGCGCTCGCAAGAGAGATGGCGGACCGGCTGCT 2008  
  
Qy 1035 GGGCAGGGAATTTATGCTCTCTGCTTCGCGCCCAACCGTTGCGCGCGGTGAAAGCGG 1094  
Db 2009 GAAGGAAGGCATCTACGTGCTGCGCTTCTGTTCCGCTGCTGCCAGGCGGCGG 2068  
  
Qy 1095 GATTGGA 1102  
Db 2069 CATCCGA 2076  
  
RESULT 13  
US-10-793-626-4459  
; Sequence 4459, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: P03480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4459  
; LENGTH: 3542  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-4459

Query Match 3.2%; Score 100.6; DB 7; Length 3542;  
Best Local Similarity 63.4%; Pred. No. 2.9e-15;  
Matches 154; Conservative 0; Mismatches 89; Indels 0; Gaps 0;  
  
Qy 2890 TGTTCGCTTTTATCAATCCATCAAAAGAAATTCGCTTTCCGAGGAGAGAGGTCAATC 2949  
Db 1 TGTTAGTTTATCAATCCAAAGAAATTCGATTCGAGTGGAGGAGTAATC 60  
  
Qy 2950 TCCGCATTCGACCCATTAGGCGCTTTACGCCGCAAACTCCATTTTTCGAGACTACT 3009  
Db 61 TACGTTTATCAACCACTCGCATTTGAAGCGGCTAATTCATTTTGTAGGAGATTACT 120  
  
Qy 3010 TACAACTCGCGGCAAGAGAGAGAGGAGGATCATATAAATGCTGAGTGTAGGCTTTG 3069  
Db 121 TAATTACAGCGGTCAACCGAATGAGGAGATTATCGCATGATTGAAGATTTAGGTTG 180  
  
Qy 3070 AAGTTGAATCAGTCGAAGAAATGAAGCTAGTTTAACTGCGAAAAGCTGAAAGAAATCAAT 3129  
Db 181 AAATCGACAGTTAAATATTAATAAATAATTTTCAATAAATGTAATATAGTTTAAAT 240  
  
Qy 3130 AAA 3132  
Db 241 AAA 243  
  
RESULT 14  
US-11-055-822-643  
; Sequence 643, Application US/11055822  
; Publication No. US20050260707A1  
; GENERAL INFORMATION:  
; APPLICANT: Pompejus, Markus  
; APPLICANT: Kroger, Burkhard  
; APPLICANT: Schroder, Hartwig

Query Match	3.0%;	Score 95;	DB 11;	Length 1125;
Best Local Similarity	47.7%;	Pred. No. 4.7e-14;		
Matches 278;	Conservative 0;	Mismatches 305;	Indels 0;	Gaps 0;
Qy	2372	GGTGAATAAGGAACGCTGCTTGAAGCGCGGAAGCGGCGCACGATCTGAATATCGGCAC	2431	
Db				
Qy	373	GCTGGATATTCGGATCTGTTGAAGCGGCTAAACAGACCGGAAAACACTGGCGGTACCGA	432	
Db				
Qy	2432	ATATTGTATCGTGGCAAGCGCGCAGAGTCCGCTCTAAACAGAGAAGTGGATCAGGTCGTAGA	2491	
Db				
Qy	433	ATTCTGTATCGTCGCGCAGTCAAGGGCGCTGATGAGAGGCTCATGACCCAGCTGGAGGA	492	
Db				
Qy	2492	TGGGGTTCAGGAANATTAAGAGAGCTATGCACTGAAGATTTTGTCATGCTCTTGGACTGTT	2551	
Db				
Qy	493	AGCAGTCTCTCGCATTCACCTGAAGTTGAAATTGAAGTCGCAGCATCGATCGGAACGTT	552	
Db				
Qy	2552	GAAGCCAGACGACGCGGAAGCGGCTCAAAGATGACGAGTAGACCGCTATATCATATATT	2611	
Db				
Qy	553	AAATAGGAACAGTGGATCGCTCGCTGCTCGCGCGTGCAACGCTCAACCAATATT	612	
Db				
Qy	2612	GAATAGCTCA CAGAGAAACCAATTCAAACATCAACACCTCA CATA CACATGACAGAGT	2671	
Db				
Qy	613	GGAAACTGCGCGTTCCCTATTTCCCTGAAGTTGTCAACCACTCATACATGGGAAGACGCCG	672	
Db				
Qy	2672	CAATACGGTTGAATTCGCAAAAGNATCGGGGCTCTCCGTGTTTCAGCGGCCATTATCGG	2731	
Db				
Qy	673	CGAACTTTTCGCGCTGTGGCAGAACTGAAATGGAAGTCTGTTCCGCGCGGAATCTTTAGG	732	
Db				
Qy	2732	GATGAAGGAGACGAAACAGGATGCA TTGACATCGCCAAAAGCTTGAAGGCTCTTTGACGC	2791	
Db				
Qy	733	AATGGGCGAACTTTAGAGCAGCGCGCGAGTTTGCCGTGCGAGCTGGCGGAGCTTTGATCC	792	
Db				
Qy	2792	GGATTCATTCCTGTGTAATTTTTTGTGATGCAATTTGATGGCACGCGCTTGAAGAGCGCGTCAA	2851	
Db				

[illegible]

Search completed: February 12, 2006, 03:17:01  
Job time : 371 secs

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame3plus.n2p.model

Run on: February 10, 2006, 02:17:17 ; Search time 92.7 Seconds

(without alignments)  
2991.757 Million cell updates/sec

Title: US-10-681-086-1

Perfect score: 5517

Sequence: 1 ggatccacagggttacgacg.....atcggtatgatgcgaattc 3156

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 4886326

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=rlp  
-Q=/abs/ABSSWEB\_spool/US10681086/runat\_09022006\_140514\_11701/app\_query.fasta.1  
-DB=A\_Geneseq -QMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02p  
-USER=US10681086 @CCN 1.1.348 @runat\_09022006\_140514\_11701 -NCPU=6 -ICPU=3  
-NO MMAP -NEG\_SCORES=0 -WAIT\_DSPBLOCK=100 -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A\_Geneseq\_21.\*

1:	Geneseqp1980s.*
2:	Geneseqp1990s.*
3:	Geneseqp2000s.*
4:	Geneseqp2001s.*
5:	Geneseqp2002s.*
6:	Geneseqp2003as.*
7:	Geneseqp2003bs.*
8:	Geneseqp2004s.*
9:	Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1836	33.3	389	8 ADS44825	Ads44825 Bacterial
2	1333	24.2	253	2 AAW31408	Aaw31408 Ant (4')-I
3	1333	24.2	253	2 AAW31402	Aaw31402 Ant (4')-I
4	1333	24.2	253	4 AAB62515	Aab62515 S. aureus
5	1333	24.2	253	4 AAB29794	Aab29794 S. aureus
6	1318	23.9	253	4 AAB62512	Aab62512 S. aureus
7	1267	23.0	253	4 AAB62513	Aab62513 Mutant ka
8	1222	22.1	253	4 AAB62514	Aab62514 Mutant ka
9	971	17.6	331	1 AAP70328	Aap70328 Biotin-sy

10	957	17.3	337	1	AAP81191	Aap81191 Biotin sy
11	942	17.1	387	2	AAW30520	Aaw30520 Kurthia s
12	932	16.9	393	1	AAP81194	Aap81194 Biotin bi
13	851	15.4	338	2	AAW30521	Aaw30521 Kurthia s
14	843	15.3	392	7	ADF41704	Adf41704 Bacillus
15	843	15.3	392	8	ADS44705	Ads44705 Bacterial
16	828	15.0	321	6	ABU43285	Abu43285 Protein e
17	828	15.0	327	5	ABP38626	Abp38626 Staphyloc
18	828	15.0	337	8	ADS05126	Ads05126 Staphyloc
19	828	15.0	332	4	AAG81723	Aag81723 S. epider
20	828	15.0	332	4	AAG82734	Aag82734 S. epider
21	809	14.7	332	6	ABU17909	Abu17909 Protein e
22	793	14.4	318	4	AAU34288	Aau34288 Staphyloc
23	793	14.4	335	4	AAU37532	Aau37532 Staphyloc
24	793	14.4	335	4	AAU37187	Aau37187 Staphyloc
25	793	14.4	335	6	ABU16098	Abu16098 Protein e
26	793	14.4	335	6	ABM73372	Abm73372 Staphyloc
27	778	14.1	321	6	ABU43752	Abu43752 Protein e
28	767	13.9	371	8	ADS44565	Ads44565 Bacterial
29	712	12.9	392	8	ADS24461	Ads24461 Bacterial
30	712	12.9	395	8	ADS28438	Ads28438 Bacterial
31	711	12.9	396	6	ABU18160	Abu18160 Protein e
32	710	12.9	388	8	ADS27616	Ads27616 Bacterial
33	705	12.8	395	4	AAU36848	Aau36848 Staphyloc
34	704	12.8	385	4	AAU34079	Aau34079 Staphyloc
35	704	12.8	396	6	ABU30082	Abu30082 Protein e
36	704	12.8	414	7	ADC96144	Adc96144 E. faeciu
37	700	12.7	393	8	ADS29218	Ads29218 Bacterial
38	699	12.7	324	6	ABU17623	Abu17623 Protein e
39	699	12.7	395	6	ABM72451	Abm72451 Staphyloc
40	699	12.7	398	6	ABU43750	Abu43750 Protein e
41	699	12.7	400	8	ADS41778	Ads41778 Bacterial
42	698.5	12.7	372	8	ADN18405	Adn18405 Bacterial
43	695.5	12.6	400	6	ADB08430	Adb08430 Alloioococ
44	692	12.5	398	4	AAB96482	Aab96482 Putative
45	692	12.5	398	8	ADS43254	Ads43254 Bacterial

#### ALIGNMENTS

RESULT 1

ADS44825  
ID ADS44825 standard; protein; 389 AA.

XX ADS44825;

XX AC

DT 02-DEC-2004 (first entry)

XX XX

DE Bacterial polypeptide #23255.

XX XX

KW Recombinant DNA construct; transformed plant; improved plant property;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
KW pathogen tolerance; pest tolerance; plant disease resistance;  
KW cell cycle pathway modification; plant growth regulator;  
KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
KW bacterial polypeptide.

OS Bacteria.

XX XX  
PN US2003233675-A1.

XX 18-DEC-2003.

XX 20-FEB-2003; 2003US-00369493.

XX 21-FEB-2002; 2002US-0360039P.

XX (CAOV/) CAO Y.

XX (HINK/) HINKLE G J.

XX (SLAT/) SLATER S C.

XX (CHEN/) CHEN X.

XX (GOLD/) GOLDMAN B S.

XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
XX WPI; 2004-061375/06.  
XX New recombinant DNA construct comprising a promoter positioned to provide  
XX for expression of a polynucleotide encoding a polypeptide from a  
XX microbial source, useful for producing plants with improved properties.  
XX Claim 1; SEQ ID NO 23255; 122pp; English.  
XX The invention relates to a recombinant DNA construct comprising a  
XX promoter functional in a plant cell, where the promoter is positioned to  
XX provide for expression of a polynucleotide encoding a polypeptide from a  
XX microbial source. The invention also relates to a transformed plant  
XX comprising the recombinant DNA construct and a method of producing a  
XX transformed plant having an improved property. The plant is a crop plant  
XX such as maize or soybean. The method of producing a transformed plant  
XX having an improved property comprises transforming a plant with the  
XX recombinant DNA construct and growing the transformed plant, where the  
XX polynucleotide or polypeptide is useful for improving plant properties.  
XX The recombinant DNA construct is useful for producing plants with  
XX improved plant properties, e.g. improved cold, heat or drought tolerance,  
XX tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
XX increased resistance to plant disease, better growth rate by modification  
XX of the cell cycle pathway with plant growth regulators, increased rate of  
XX homologous recombination, modified seed oil or protein yield and/or  
XX content, improved yield by modification of carbohydrate, nitrogen or  
XX phosphorus use and/or uptake, by modification of photosynthesis or by  
XX providing improved plant growth and development under at least one stress  
XX condition, improved lignin production or improved galactomannan  
XX production. This sequence represents a bacterial polypeptide used in the  
XX scope of the invention. Note: The sequence data for this patent did not  
XX form part of the printed specification but was obtained in electronic  
XX format from USPTO at seqdata.uspto.gov/sequence.html.  
XX SQ Sequence 389 AA;

Alignment Scores:  
Pred. No.: 4,98e-185 Length: 389  
Score: 1836.00 Matches: 359  
Percent Similarity: 100.0% Conservative: 1  
Best Local Similarity: 99.7% Mismatches: 0  
Query Match: 33.3% Indels: 0  
DB: 8 Gaps: 0

US-10-681-086-1 (1-3156) x ADS44825 (1-389)

QY 22 TTGAAGATTGATTCTGGTTAAACGAGCGGTTAGACAGAAATGAAAGCCGGGTACAT 81  
DB 1 MetLysIleAspSerTrpLeuAsnGluArgLeuAspArgMetLysGluAlaGlyValHis 20  
QY 82 CGTAACCTCGCGTCAATCGATGGAGCGCGGTTCCAGAGAGAAATATTGATGGCGAAAT 141  
DB 21 ArgAsnLeuArgSerMetAspGlyAlaProValProGluArgAsnIleAspGlyGluAsn 40  
QY 142 CAACCGTCTCGTCTCAACAAATATTATGGCTCGCAACGATAGACGTTTGTATCGAT 201  
DB 41 GlnThrValTrpSerSerAsnAsnTyrLeuGlyLeuAlaSerAspArgArgLeuIleAsp 60  
QY 202 GCAGCCCAACAGCAATTCAGCAATTTGGGACAGGACGAGCGGTTACGTTTAAACGACA 261  
DB 61 AlaAlaGlnThrAlaLeuGlnGlnPheGlyThrGlySerSerGlySerArgLeuThrThr 80  
QY 262 GGCAATTCGGTCTGGCATGAAAGCTAGAAAGAGATTGCCAGCTTTAAACTGACAGAA 321  
DB 81 GlyAsnSerValTrpHisGluLysLeuGluLysLysIleAlaSerPheLysLeuThrGlu 100  
QY 322 CGGCGCTCGTCTTTTCAGCGGGTTACTTGGCCAAATGTCGGTGTCTTTCATCTCTGCCA 381  
DB 101 AlaAlaLeuLeuPheSerSerGlyTyrLeuAlaAsnValGlyValLeuSerSerLeuPro 120  
QY 382 GAAAGGAAGATGTCTTTAAGTGACCGACTCAATCATGCAAGTATGATCGACGGCTGC 441

DB 121 GluLysGluAspValIleLeuSerAspGlnLeuAsnHisAlaSerMetIleAspGlyCys 140  
QY 442 CGACTTTCTAAGCTGATACAGTTGTTATCGGCATATTGATATGATGATCTTGAAGAAC 501  
DB 141 ArgLeuSerLysAlaAspThrValValTyrArgHisIleAspMetAsnAspLeuGluAsn 160  
QY 502 AAGCTGAATCAACACACAGCGTTATCAGCGCGTATTTATCGTAACAGACGGAGTATTCAGC 561  
DB 161 LysLeuAsnGluThrGlnArgTyrGlnArgPheIleValThrAspGlyValPheSer 180  
QY 562 ATGATATCGCAATTCGCGCCCTCTTGTATCAGATCATCTCATTGCGAAACGCTATCATGCC 621  
DB 181 MetAspGlyThrIleAlaProLeuAspGlnIleSerLeuAlaLysArgTyrHisAla 200  
QY 622 TTCGTGTGCTGTGATGATGCCCAACAGAGAGTTTGGCGGATTTGGGACAGGAACG 681  
DB 201 PheValValAspAspAlaHisAlaThrGlyValLeuGlyAspSerGlyGlnGlyThr 220  
QY 682 AGTGAATACTTGGTGTGTTGTCGCGACATTTGTCGACCTTAAAGCAAGCTGTTGCG 741  
DB 221 SerGluTyrPheGlyValCysProAspIleValIleGlyThrLeuSerLysAlaValGly 240  
QY 742 GCGAAGAGAGTTTGGCGGAGATCAGCGGTCTTTCATCGACTTTTGTCTGAACCATGCC 801  
DB 241 AlaGluGlyGlyPheAlaAlaGlySerAlaValPheIleAspPheLeuLeuAsnHisAla 260  
QY 802 AGAATATTATTTTCAACCGCTATTCGCGCAGCGAGTGTGCGGTGCTCAGAGGCT 861  
DB 261 ArgThrPheIlePheGlnThrAlaIleProProAlaSerCysAlaAlaHisGluAla 280  
QY 862 TTCACATCATTTGAGCCAGCAGCGGAAACGACAGCTTTTATTTCTTATATCAGCATG 921  
DB 281 PheAsnIleIleGluAlaSerArgLysArgGlnLeuLeuPheSerTyrIleSerMet 300  
QY 922 ATCAGAACCATGCTCTGAAGAATATGGGTATGTGTGTAAGAGAGATCACACCCGATTAT 981  
DB 301 IleArgThrSerLeuLysAsnMetGlyTyrValValLysGlyAspHisThrProIleIle 320  
QY 982 CTGTAGTCAATTGGCGATGCCCATAAACCGGTCTCTATTGCTGAAAACTCGAGGGCAAG 1041  
DB 321 ProValValIleGlyAspAlaHisLysThrValLeuPheAlaGluLysLeuGlnGlyLys 340  
QY 1042 GGAATTTATGCTCTGCGCATTCGCGCCCAACCGTTCGCGCGGTGAAAGCCGATTCCGA 1101  
DB 341 GlyIleTyrAlaProAlaIleArgProThrValAlaProGlyGluSerArgIleArg 360  
RESULT 2  
AAW31408  
ID AAW31408 standard; protein; 253 AA.  
XX  
AC AAW31408;  
XX  
DT 11-MAY-1998 (first entry)  
XX  
XX Ant(4')-Ia protein.  
XX  
XX Kanamycin resistance gene; adenylyl 4'-nucleotidyltransferase type Ia;  
XX ant(4')-Ia gene; composite nucleic acid molecule synthesis;  
XX site specific DNA fragment ligation; PCR product ligation.  
OS Synthetic.  
OS Escherichia coli.  
XX  
XX WO9742330-A1.  
FN  
PD 13-NOV-1997.  
XX  
XX 06-MAY-1997; 97WO-US0007698.  
PF  
XX 06-MAY-1996; 96US-00642045.  
PR  
XX (APOL-) APOLLON INC.  
PA



```
XX      Sequence 253 AA;
SQ
Alignment Scores:
Pred. No.:      1,08e-131      Length:      253
Score:          1333.00      Matches:      252
Percent Similarity: 100.0%      Conservative: 1
Best Local Similarity: 99.6%      Mismatches: 0
Query Match:    24.2%      Indels: 0
DB:             2      Gaps: 0

US-10-681-086-1 (1-3156) x AAB62515 (1-253)
Qy 1579 GTGAATGCACCAATAATAATGACTAGACAGAAAGAATGAAGATTGTTTCATGAATAAAG 1638
Db 1 MetAsnGlyProIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
Qy 1639 GAACGATATATGGATAATATGGGATGATGTTAAGGCTATTGGTGTGTTATGGCTCTCTT 1698
Db 21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40
Qy 1699 GGTGCTCAGACTGATGGCCCTATTTCGGATATTGAGATGATGTGTGTCATGTCAACAGAG 1758
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60
Qy 1759 GAAGCAGAGTTTCAGCCATGAATGGAACACCGGTGAGTGGAAAGGTGGAAGTGAATTTTGCAT 1818
Db 61 GluAlaGluPheSerHisGluTrpThrGlyGluTrpLysValGluValAsnPheAsp 80
Qy 1819 AGCGAAGAGATTCTACTAGATTATGCAATCTCAGGTGGAATCAGATTGGCCGCTTACACAT 1878
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpProLeuThrHis 100
Qy 1879 GGTCAATTTTCTCTATTGTCGCGATTTATGATTCAGGTGATATCTAGACAAAGTGTAT 1938
Db 101 GlyGlnPhePheSerIleLeuProIleTyrAspSerGlyTyrLeuGluLysValTyr 120
Qy 1939 CAACTGCTTAATCGTGAAGACCCAAACGTTTCACAGATGCGATTGTCGCCCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAlaLeuIleVal 140
Qy 1999 GAAGAGCTGTTTGAATATGACGGCAAAATGGCGTAAATATTCGTGTCGAAGCCGCAACA 2058
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
Qy 2059 TTTCTACCATCTTGACCTGTACAGGTACAGTACCAATGGCAGTGCATGTTGATGCTGTCAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
Qy 2119 CATGCACTGTTATACACGAGCGCTTCGGTCTTAACCTGAAGCAGTTTAAGCAATCAGAT 2178
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200
Qy 2179 CTTCCTCAGGTTATGACCATCTGTGCGAGTTTCGTAATGTCGTGTCACACTTCCGACTCT 2238
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
Qy 2239 GAGAACTTCGGAATCCCTAGAGAAATTCGGAATGGGATTCAGGATGGACGAGAACGA 2298
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
Qy 2299 CACGGATATATAGTGGATGTGTCAAACGCGCATACCATTT 2337
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 4
AAB62515
ID AAB62515 standard; protein; 253 AA.
XX
AC AAB62515;
XX
DT 23-JUL-2001 (first entry)
XX
```

```
DE
XX
KW
XX
OS
XX
PN
XX
PD
XX
PF
XX
PR
XX
PA
XX
PI
XX
DR
XX
PT
XX
PS
XX
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
SQ
Alignment Scores:
Pred. No.:      1,08e-131      Length:      253
Score:          1333.00      Matches:      252
Percent Similarity: 100.0%      Conservative: 1
Best Local Similarity: 99.6%      Mismatches: 0
Query Match:    24.2%      Indels: 0
DB:             4      Gaps: 0

US-10-681-086-1 (1-3156) x AAB62515 (1-253)
Qy 1579 GTGAATGCACCAATAATAATGACTAGACAGAAAGAATGAAGATTGTTTCATGAATAAAG 1638
Db 1 MetAsnGlyProIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
Qy 1639 GAACGATATATGGATAATATGGGATGATGTTAAGGCTATTGGTGTGTTATGGCTCTCTT 1698
Db 21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40
Qy 1699 GGTGCTCAGACTGATGGCCCTATTTCGGATATTGAGATGATGTGTGTCATGTCAACAGAG 1758
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60
Qy 1759 GAACGAGAGTTTCAGCCATGAATGGAACACCGGTGAGTGGAAAGGTGGAAGTGAATTTTGCAT 1818
Db 61 GluAlaGluPheSerHisGluTrpThrGlyGluTrpLysValGluValAsnPheAsp 80
Qy 1819 AGCGAAGAGATTCTACTAGATTATGCAATCTCAGGTGGAATCAGATTGGCCGCTTACACAT 1878
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpProLeuThrHis 100
Qy 1879 GGTCAATTTTCTCTATTGTCGCGATTTATGATTCAGGTGATATCTAGACAAAGTGTAT 1938
Db 101 GlyGlnPhePheSerIleLeuProIleTyrAspSerGlyTyrLeuGluLysValTyr 120
Qy 1939 CAACTGCTTAATCGTGAAGACCCAAACGTTTCACAGATGCGATTGTCGCCCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAlaLeuIleVal 140
Qy 1999 GAAGAGCTGTTTGAATATGACGGCAAAATGGCGTAAATATTCGTGTCGAAGCCGCAACA 2058
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
Qy 2059 TTTCTACCATCTTGACCTGTACAGGTACAGTACCAATGGCAGTGCATGTTGATGCTGTCAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
Qy 2119 CATGCACTGTTATACACGAGCGCTTCGGTCTTAACCTGAAGCAGTTTAAGCAATCAGAT 2178
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200
Qy 2179 CTTCCTCAGGTTATGACCATCTGTGCGAGTTTCGTAATGTCGTGTCACACTTCCGACTCT 2238
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
Qy 2239 GAGAACTTCGGAATCCCTAGAGAAATTCGGAATGGGATTCAGGATGGACGAGAACGA 2298
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
Qy 2299 CACGGATATATAGTGGATGTGTCAAACGCGCATACCATTT 2337
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

S. aureus wild-type kanamycin nucleotidyltransferase (KNT) .
Kanamycin nucleotidyltransferase; KNT; thermostability; mutation; mutant; marker; thermophilic; Thermus thermophilus.
Staphylococcus aureus.
EPI097990-A2.
09-MAY-2001.
27-OCT-2000; 2000EP-00309475.
29-OCT-1999; 99JP-00309616.
(RIKE ) RIKEN KK.
Yokoyama S, Hoseki J, Yano T, Koyama Y, Kuramitsu S; Kagamiyama H;
WPI; 2001-337112/36.
N-PSDB; AAF83650.
Novel mutant kanamycin nucleotidyltransferase useful for screening thermophilic bacteria such as Thermus thermophilus, has point mutations at specific positions and improved thermostability.
Disclosure; Page 22-24; 30pp; English.
The invention relates to mutant kanamycin nucleotidyltransferases (KNT) having improved thermostability. The mutant has one or more point mutations selected from Met57Leu, Ala62Val, Ser94Pro, Ser203Pro, Asp205Val, His207Gln, Ser220Pro, Ile234Val and Thr238Ala as against a WT-KNT (AAB62515). A KNT gene encoding the mutant can be used as a selective marker for thermophilic bacteria. The marker can be used for screening thermophilic bacteria, preferably Thermus thermophilus. The present sequence represents a S. aureus wild-type KNT protein. The WT-KNT protein from which the mutants are constructed differs from this protein at position 80 and position 130
```



Db 101 GlyGlnPheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120  
QY 1939 CAAACTGCTAAATCGGTAGAGCCCAACGTTCCACGATCGGATTGTGCGCTTATCGTA 1998  
Db 121 GlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAlaLeuIleVal 140  
QY 1999 GAAGAGCTGTTGAATATGATGAGCAAAATGCGTAATATTCGTGTGCAAGGACCGACAACA 2058  
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160  
QY 2059 TTCTACCATCCTGACTGACGTAGAGTAGCAATGCGAGTGCCATGTTGATGTTGTCAT 2118  
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180  
QY 2119 CATCGCATCTGTATACGACGCGCTGCTGCTTACTCAAGCAGTTAAGCAATCAGAT 2178  
Db 181 HisArgIleCysIleThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200  
QY 2179 CTTCTCTCAGGTTATGACCATCTGTGTCAGTTCGTATGCTGTCACACTTTCGCACTCT 2238  
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220  
QY 2239 GADAACTTCGAAATCGCTAGAGAAATTTCTGGAATCGGATTCGAGTGCAGACGACGA 2298  
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240  
QY 2299 CACGGATATATAGTGGATGTGCAAAACGCATACCATTT 2337  
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 5

AAB29794  
ID AAB29794 standard; protein; 253 AA.  
XX AAB29794;  
AC AAB29794;  
XX  
DT 09-MAR-2001 (first entry)  
XX  
DE S. aureus adenyl 4'-nucleotidyltransferase type Ia (ant(4')-Ia).  
XX  
KW Composite nucleic acid molecule synthesis; ligation; bridge oligomer;  
KW bridging oligonucleotide; bridge oligomer; heat-stable ligase;  
KW chimeric kanamycin resistance gene; aph(3')-Ia promoter; terminator;  
KW aminoglycoside 3'-phosphotransferase type Ia; modified ant(4')-Ia;  
KW adenyl 4'-nucleotidyltransferase; gentamicin-sensitive;  
KW selectable marker; clinical expression vector.  
XX  
OS Staphylococcus aureus.  
XX  
XX US6143527-A.  
XX  
XX 07-NOV-2000.  
XX  
XX 06-MAY-1997; 97US-00852268.  
XX  
XX 06-MAY-1996; 96US-00642045.  
XX  
XX (AMHP ) AMERICAN HOME PROD CORP.  
XX  
XX Samuel M, Satishchandran C, Pachuk C;.  
XX  
XX WPI; 2001-006434/01.  
XX  
XX N-PSDB; AAC81670.  
XX  
XX Producing a composite nucleic acid molecule for site specific ligation of  
XX digested DNA fragments, involves ligation of digested double-stranded DNA  
XX molecules by DNA ligases and bridging oligonucleotides.  
XX  
XX Example 1; Col 25-28; 41pp; English.  
XX  
XX The invention relates to a novel method of synthesizing a composite  
XX nucleic acid molecule. The method involves denaturing two separate double

CC stranded nucleic acids (N1 and N2), and hybridising them to a bridging  
CC oligonucleotide comprising, in the 5' to 3' direction, a 10-40 nucleotide  
CC sequence complementary to the 5' end of the first strand of N1 and a 10-  
CC 40 nucleotide sequence complementary to the 3' end of the first strand of  
CC N2. The N1 first strand 5' end and the N2 first strand 3' end are thus  
CC oriented adjacent to each other, and a heat-stable ligase is then used to  
CC ligate the N1 and N2 first strands to form the first strand of the  
CC composite nucleic acid. The composite first strand is then denatured from  
CC the bridging oligonucleotide, and allowed to reanneal to the  
CC complementary strands of N1 and N2. The complementary strands are then  
CC ligated to form a composite complementary strand. The invention also  
CC encompasses the ligation of a number of nucleic acid molecules by the  
CC method, so that the mixture comprises a number of nucleic acid molecules  
CC and a number of bridging oligonucleotides. The method is useful for the  
CC site-specific ligation of DNA fragments generated by PCR, restriction  
CC enzyme digestion, DNase digestion, chemical cleavage, and enzymatic or  
CC chemical synthesis. Large nucleic acid molecules are generated from  
CC smaller ones without the need to treat the ends to ensure specific  
CC binding of two molecules in correct orientation. The bridge oligomer  
CC template ensures correct orientation and specific ligation. Ligation of  
CC multiple components can take place without the need for intervening  
CC steps, using a heat-stable ligase and the necessary starting materials.  
CC In an exemplification of the invention, a chimeric kanamycin resistance  
CC gene was synthesised using the method of the invention. The chimeric  
CC kanamycin gene comprises the promoter and terminator of the Escherichia  
CC coli kanamycin resistance gene aph(3')-Ia (which encodes aminoglycoside  
CC 3'-phosphotransferase type Ia) and the open reading frame of the  
CC Staphylococcus aureus kanamycin resistance gene ant(4')-Ia (which encodes  
CC adenyl 4'-nucleotidyltransferase type Ia). The ant(4')-Ia open reading  
CC frame was also modified such that the wild-type gram-positive GTG  
CC translation initiation codon was altered to the gram-negative preferred  
CC ATG initiation codon, and a nearby out-of-frame ATG codon was altered to  
CC ACG. Neither of these alterations alter the amino acid sequence of the  
CC protein as expressed in S. aureus. The chimeric kanamycin resistance gene  
CC has the activity of the S. aureus ant(4')-Ia protein (including being  
CC sensitive to gentamicin) but is able to be expressed in E. coli. The  
CC narrower range of activity of the chimeric kanamycin resistance gene  
CC against aminoglycoside antibiotics makes it particularly useful as a  
CC selectable marker in the production of vectors for clinical use in  
CC humans. The present sequence represents the amino acid sequence of the  
CC Staphylococcus aureus ant(4')-Ia kanamycin resistance protein, which is  
CC encoded in an Escherichia coli-expressible form by the chimeric kanamycin  
CC resistance gene produced using the method of the invention  
XX  
SQ Sequence 253 AA;  
Alignment Scores:  
Pred. No.: 1.08e-131 Length: 253  
Score: 1333.00 Matches: 252  
Percent Similarity: 100.0% Conservative: 1  
Best Local Similarity: 99.6% Mismatches: 0  
Query Match: 24.2% Indels: 0  
DB: 4 Gaps: 0  
US-10-681-086-1 (1-3156) x AAB29794 (1-253)  
QY 1579 GTCAATGACCAATAATATGACTAGAGAGAAAGAAATGATTTTCATGAATTAAG 1638  
Db 1 MetasnglyprolilelleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20  
QY 1639 GAACGAATATGGATAAATATGGGATGATTTTAAGGCTATTGGTGTATGGCTCTCTT 1698  
Db 21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40  
QY 1699 GGTGTCAGACTGATGGGCCCTATTCCGATATTGAGATGATGTCATGTCACACAGAG 1758  
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60  
QY 1759 GAAGCAGAGTTTCAGCCATGAATCGACCAACCGGTAGTGGAGGTGAAGTGAATTTTGTAT 1818  
Db 61 GluAlaGluPheSerHisGluTrpThrThrGlyGluTrpLysValGluValAsnPheAsp 80  
QY 1819 AGCGAAGAGATTCTACTAGATTATGCAATTCAGGTGGAATCAGATTGGCCGCTTACACAT 1878

```
|||||
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTyrProLeuThrHis 100
QY 1879 GGTCAATTTTCTCTATTTCCCGATTTTATGATTCAGGTGGTACTTAGAGAAAGTGTAT 1938
Db 101 GlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120
QY 1939 CAAATCGTAAATCGTAGAAGCCCAAAACGTTCCACAGATGGATTTGTGCCCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAlaLeuIleVal 140
QY 1999 GAAGAGCTGTTGAATATGCGAGGCAAAATGGCGTAATATTCGTGTGCAAGGACCGACACA 2058
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
QY 2059 TTTCTACCATCTTGACCTACAGTACAGTACAGTGGCAGGTGCATGTTGATGGTCTGCAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
QY 2119 CATCGCATCTGTATACGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGAT 2178
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200
QY 2179 CTTCTCTCAGGTATGACCATCTGTCGAGTTCGTAATGTCTGGTCAACTTTCGACACTCT 2238
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
QY 2239 GAGAACTCTCGAATCGCTAGAGAAATTCGGAATGGATTCAGGATGCGACGACGACGACGA 2298
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
QY 2299 CACGATATATAGTGGATGTGTCAAAACGACATACCATTT 2337
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 6
AAB62512
ID AAB62512 standard; protein; 253 AA.
XX
AC AAB62512;
XX
DT 23-JUL-2001 (first entry)
XX
DE S. aureus kanamycin nucleotidyltransferase WT*KNT protein.
XX
KW Kanamycin nucleotidyltransferase; KNT; thermostability; mutation; mutant;
KW marker; thermophilic; thermus thermophilus.
XX
OS Staphylococcus aureus.
XX
FH Key Location/Qualifiers
FT Misc-difference 80
FT /label= D80Y
FT /note= "wild-type Asp is replaced with Tyr"
FT Misc-difference 130
FT FT T130K
FT /label= T130K
FT /note= "wild-type Thr is replaced with Lys"
XX
FN EP1097990-A2.
XX
XX 09-MAY-2001.
XX
XX 27-OCT-2000; 2000EP-00309475.
XX
XX 29-OCT-1999; 99JP-00309616.
XX
XX (RIKE ) RIKEN KK.
XX
XX Yokoyama S, Hoseki J, Yano T, Koyama Y, Kuramitsu S;
XX Kagamiyama H;
XX WPI; 2001-337112/36.
XX
```

```
PT Novel mutant kanamycin nucleotidyltransferase useful for screening
PT thermophilic bacteria such as Thermus thermophilus, has point mutations
XX at specific positions and improved thermostability.
PS Claim 1; Page 12-13; 30pp; English.
XX
CC The invention relates to mutant kanamycin nucleotidyltransferases (KNT)
CC having improved thermostability. The mutant has one or more point
CC mutations selected from Met57Leu, Ala62Val, Ser94Pro, Ser203Pro,
CC Asp206Val, His207Gln, Ser220Pro, Ile234Val and Thr238Ala as against a
CC WT*KNT (AAB62512). A KNT gene encoding the mutant can be used as a
CC selective marker for thermophilic bacteria. The marker can be used for
CC screening thermophilic bacteria, preferably Thermus thermophilus. The
CC present sequence represents the S. aureus WT*KNT protein from which the
CC mutants are constructed. This differs from the wild-type KNT protein
CC (AAB62515) at position 80 and position 130
XX
SQ Sequence 253 AA;

Alignment Scores:
Pred. No.: 4.26e-130 Length: 253
Score: 1318.00 Matches: 250
Percent Similarity: 99.2% Conservative: 1
Best Local Similarity: 98.8% Mismatches: 2
Query Match: 23.9% Indels: 0
DB: Gaps: 4

US-10-681-086-1 (1-3156) x AAB62512 (1-253)
QY 1579 GTGAATCGACAATAATAATGACTAGAGAGAAAGATGAAGATGTTTCATGAATTAAG 1638
Db 1 MetAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
QY 1639 GAACGATATTTGGATAAATATGGGATGATGTTAAGGCTATTGCTGTTATGGCTCTCT 1698
Db 21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40
QY 1699 GGTCTGTCAGACTGATGGGCCCTTATTCGATATTCAGATGATGTCATGTCATCAACAGAG 1758
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60
QY 1759 GAACGAGCTTCAGCCCATGAATGAGCAACCGGTGAGTGGAGGTGGAAGTGAATTTGAT 1818
Db 61 GluAlaGluPheSerHisGluTrpThrThrGlyLysValGluValAsnPheTyr 80
QY 1819 AGCGAAGAGATTCCTACTAGATATTGTCATCTCAGGTGGAATCAGATTCGCCCTTACACAT 1878
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTyrProLeuThrHis 100
QY 1879 GGTCAATTTTCTCTATTTGCCGATTTATGATTCAGGTGGATACCTTAGAGAAAAGTGTAT 1938
Db 101 GlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120
QY 1939 CAAATCGTAAATCGTAGAAGCCCAAAACGTTCCACAGTGGATTTGTGCCCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140
QY 1999 GAAGAGCTGTTGAATATGCGAGGCAAAATGGCGTAATATTCGTGTGCAAGGACCGACACA 2058
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
QY 2059 TTTCTACCATCTTGACCTACAGTACAGTACAGTGGCAGGTGCATGTTGATGGTCTGCAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
QY 2119 CATCGCATCTGTATACGACGAGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGAT 2178
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200
QY 2179 CTTCTCTCAGGTATGACCATCTGTCGAGTTCGTAATGTCTGGTCAACTTTCGACACTCT 2238
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
```

QY 2239 GAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTGAGGAGTCGACAGACGA 2298  
 DB 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240  
 QY 2299 CACGGATATAGTGGATGCTCAAAACGATACCATTT 2337  
 DB 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253  
 RESULT 7  
 AAB62513  
 ID AAB62513 standard; protein; 253 AA.  
 KW  
 AC AAB62513;  
 XX  
 DT 23-JUL-2001 (first entry)  
 XX  
 DE Mutant kanamycin nucleotidyltransferase KT3-11.  
 XX  
 KW Kanamycin nucleotidyltransferase; KNT; thermostability; mutation; mutant;  
 KW marker; thermophilic; Thermus thermophilus.  
 XX  
 OS Staphylococcus aureus.  
 OS Synthetic.  
 XX  
 PN EP1097990-A2.  
 XX  
 XX 09-MAY-2001.  
 XX  
 PF 27-OCT-2000; 2000EP-00309475.  
 XX  
 PR 29-OCT-1999; 99JP-00309616.  
 XX  
 PA (RIKE ) RIKEN KK.  
 XX  
 PI Yokoyama S, Hoseki J, Yano T, Koyama Y, Kuramitsu S;  
 PI Kagamiyama H;  
 XX  
 DR WPI; 2001-337112/36.  
 XX  
 XX  
 PT Novel mutant kanamycin nucleotidyltransferase useful for screening  
 PT thermophilic bacteria such as Thermus thermophilus, has point mutations  
 PT at specific positions and improved thermostability.  
 PS  
 PS Claim 2; Page 14-15; 30pp; English.  
 XX  
 CC The invention relates to mutant kanamycin nucleotidyltransferases (KNT)  
 CC having improved thermostability. The mutant has one or more point  
 CC mutations selected from Met57Leu, Ala62Val, Ser94Pro, Ser203Pro,  
 CC Asp206Val, His207Gln, Ser220Pro, Ile234Val and Thr238Ala as against a  
 CC WT-KNT (AAB62512). A KNT gene encoding the mutant can be used as a  
 CC selective marker for thermophilic bacteria. The marker can be used for  
 CC screening thermophilic bacteria, preferably Thermus thermophilus. The  
 CC present sequence represents a S. aureus mutant KNT protein KT3-11  
 XX  
 SQ Sequence 253 AA;  
 Alignment Scores:  
 Pred. No.: 1,12e-124 Length: 253  
 Score: 1267.00 Matches: 240  
 Percent Similarity: 96.4% Conservative: 4  
 Best Local Similarity: 94.9% Mismatches: 9  
 Query Match: 23.0% Indels: 0  
 DB: 4 Gaps: 0  
 US-10-681-086-1 (1-3156) x AAB62513 (1-253)  
 QY 1579 GTCAATGCAACAATAATGACTAGAGAAAGATGAAGATTGTTCTATGAAATTAAG 1638  
 DB 1 MetLysGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20  
 QY 1639 GAACGAATATGGATAAATGAGGATGATGTTAAGCGCTATTGGTGTATTGGCTCTCTT 1698  
 DB 21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40

QY 1699 GGTGCTCAGACTGATGGGCCCTATTTCGGATATTGAGATGATGTGTGTCAACAGAG 1758  
 DB 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60  
 QY 1759 GAAGCAGAGTTGAGCCATGAATGGACACCGGTGAGTGGAGGTGGAAGTGAATTTGAT 1818  
 DB 61 GlyAlaGluPheSerTyrGluTrpThrThrGlyGluTrpLysAlaGluValAsnPheTyr 80  
 QY 1819 AGCGAAGAGATTCTACTAGATTATGATCTCAGGTGGAATCAGATTGGCGCTTACACAT 1878  
 DB 81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluSerAspTrpProLeuThrHis 100  
 QY 1879 GGTCAATTTTCTCTATTTCGCGATTATTGATTCAGGTGGATCTTAGAGAAAGTGTAT 1938  
 DB 101 GlyArgPheSerIleLeuProIleTyrAspProGlyTyrPheGluLysValTyr 120  
 QY 1939 CAAACTGCTAAATCGGTAGAGCCCAACGGTTCACGATCGGATTTGTGCCCTTATCGTA 1998  
 DB 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140  
 QY 1999 GAAGAGCTGTTGAAATATGACGAGCAAAATGGCGTAATATTTCGTGCAAGGACGACACA 2058  
 DB 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160  
 QY 2059 TTTTCTACCATCTTGTGACTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTTGCTGCT 2118  
 DB 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180  
 QY 2119 CATCGCATCTGTTATACGACGCGCTTCGGTCTTAACCTGAAGCAGTGAAGCAATCAGAT 2178  
 DB 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp 200  
 QY 2179 CTTCTCTCAGGTTATGACCATCTGTCAGTTCGTAATGTCGTCAACTTTCGACTCT 2238  
 DB 201 LeuProSerGlyTyrAspHisLeuCysGlnLeuValMetSerGlyGlnLeuSerAspSer 220  
 QY 2239 GAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTGAGGAGTCGACAGACGA 2298  
 DB 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240  
 QY 2299 CACGGATATAGTGGATGCTCAAAACGATACCATTT 2337  
 DB 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253  
 RESULT 8  
 AAB62514  
 ID AAB62514 standard; protein; 253 AA.  
 XX  
 AC AAB62514;  
 XX  
 DT 23-JUL-2001 (first entry)  
 XX  
 DE Mutant kanamycin nucleotidyltransferase HTK.  
 XX  
 KW Kanamycin nucleotidyltransferase; KNT; thermostability; mutation; mutant;  
 KW marker; thermophilic; Thermus thermophilus.  
 XX  
 OS Staphylococcus aureus.  
 OS Synthetic.  
 XX  
 PN EP1097990-A2.  
 XX  
 XX 09-MAY-2001.  
 XX  
 PF 27-OCT-2000; 2000EP-00309475.  
 XX  
 PR 29-OCT-1999; 99JP-00309616.  
 XX  
 PA (RIKE ) RIKEN KK.  
 XX  
 PI Yokoyama S, Hoseki J, Yano T, Koyama Y, Kuramitsu S;  
 PI Kagamiyama H;



```
QY 2373 GTGAATAAGGAACGCTGCTTGAAGCGCGAAGCGCGCACGATCTGAATATCGGCACA 2432
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 86 IleThrLysGluGluIleLeuAlaGlyAlaLysArgAlaPheGluAsnLysIleGlyThr 105

QY 2433 TATTGATCTGCGACGCGCAGAGTCCGCTCTAAACAGAGAAGTGATGAGTCTGATAGAT 2492
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 106 TyrCysIleValAlaSerGlyArgGlyProThrArgLysAspValAsnValSerGlu 125

QY 2493 GCGGTTTCAGGAATTAAGAGACGCTGAGCTGAAGATTGTGCATGCTCTTGACACTGTTG 2552
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 126 AlaValGluGluIleLysAlaLysTyrGlyLeuLysValCysAlaCysLeuGlyLeuLeu 145

QY 2553 AAGCCACAGCAGCGCGAAGCGCTCAAGATGCGAGGTAGACCGCTATATCATATTTG 2612
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 146 LysGluGluGlnAlaGlnLeuLysGluAlaGlyValAspArgTyrAsnHisAsnLeu 165

QY 2613 AATAGCTCAGAGAAACCACTCAACATCAACACCTCACAATACATACATGATGACAGTCT 2672
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 166 AsnThrSerGluArgHisHisSerTyrIleThrThrHisThrTyrGluAspArgVal 185

QY 2673 AATACGTTCAATTCGCAAGAAATCGGGCTGCTCTCGGTTCAGCGCCATTATCGGG 2732
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 186 AsnThrValGluValValLysHisGlyIleSerProCysSerGlyAlaIleIleGly 205

QY 2733 ATGAAGGAGACGAACACAGGATGTCATTGACATCGCCAAAGCTTGAAGGCTCTTGACGCG 2792
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 206 MetLysGluThrLysMetAspValValGluIleAlaArgAlaLeuHisGlnLeuAspAla 225

QY 2793 GATTCCATTCCTGTGAATTTTTCATGCAATTGATGGCAGCGCTTAGAAGCGGTCAAC 2852
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 226 AspSerIleProValAsnPheLeuHisAlaIleAspGlyThrLysLeuGluGlyThrGln 245

QY 2853 GAATTAACCCGCTGATTGTTTAAAGTCTCGCGCTGTTCCGTTTATCAATCCATCA 2912
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 246 AspLeuAsnProArgTyrCysLeuLysValLeuAlaLeuPheArgTyrMetAsnProser 265

QY 2913 AAAGAATTCGCATTTCCGAGGAAGAGAGTCAATCTCCGACATTCGACCATTAGGG 2972
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 266 LysGluIleArgIleSerGlyGlyArgGluValAsnLeuGlyPheLeuGlnProPheGly 285

QY 2973 CTTTACGCCCAACTCCATTTTTCGCGAGACTACTTAAACAACATCGCCGGGCAAGAGGAG 3032
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 286 LeuTyrAlaAlaAsnSerIlePheValGlyAspTyrLeuThrThrGluGlyGlnGluAla 305

QY 3033 ACGGAGGATCAATAAATGCTGAGTATTAGGCTTTGAAGTTGAATCAGTCAAGAAATG 3092
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 306 AsnSerAspTyrArgMetLeuGluAspLeuGlyPheGluIleGluLeuThrGlnLysGln 325

QY 3093 AAG 3095
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 326 Glu 326

RESULT 10
ID AAP81191
AC standard; protein; 337 AA.
XX AAP81191;
XX AAP81191;
XX
DT 21-OCT-2004 (revised)
DT 25-MAR-2003 (revised)
DT 04-DEC-1990 (first entry)
XX
DE Biotin synthetase (Bio B) of B.sphaericus IFO 3525.
XX
KW biotin biosynthesis; fermentative biotin production; plasmid pTG1400;
KW biotin synthetase; Bio B; seborrhoeic dermatitis treatment.
XX
OS Bacillus sphaericus.
OS Unidentified.
XX
PN EP266240-A.
XX
```

```
PD 04-MAY-1988.
XX
XX 28-SEP-1987; 87BP-00402157.
XX
PR 30-SEP-1986; 86FR-00013603.
PR 18-MAY-1987; 87FR-00006916.
XX
PA (TRGE ) TRANSGENE SA.
XX
PI Gloeckler, Speck D, Lemoine Y;
XX
XX WPI; 1988-121306/18.
DR N-PSDB; AAN81546.
XX
XX New DNA sequences encoding enzymes involved in biotin biosynthesis -
PT isolated from Bacillus sphaericus, and transformants useful in
PT fermentative biotin prodn.
XX
PS Disclosure; Page: 5lpp; French.
XX
CC DNA was isolated from B.sphaericus IFO 3525, cut with HindIII and the
CC fragments sub-cloned into pBR322. The recombinant plasmids were used to
CC transform E.coli bio- mutants. Plasmid pTG1400 was present in one clone
CC (E.coli C600; CNCM I-608) which showed complementation for bio A and bio
CC D. The plasmid contains a 4.3kb insert which contains 4 long open reading
CC frames (LORFs) including that which encodes Bio B. This enzyme is well
CC characterised. A second clone was isolated which contained DNA encoding
CC bio F and bio C. The insert present in the plasmid of this clone and the
CC 4.3kb insert from pTG1400 were cloned together in pBR322 to produce a
CC plasmid (pTG1440) which complements biotin auxotrophy in a bio delta FCD
CC mutant. Transformed cells are used to produce biotin by culture in a
CC medium contg pimelic acid and/or biotin vitamer. Co-culture of two
CC transformants is preferred where one synthesises the vitamer from pimelic
CC acid and the other converts the vitamer to biotin. See also AAN81551.
CC (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to
CC correct PI field.)
CC
CC Revised record issued on 21-OCT-2004 : Correction to location
XX
XX Sequence 337 AA;
SQ
Alignment Scores:
Pred. No.: 1,19e-91 Length: 337
Score: 957.00 Matches: 180
Percent Similarity: 84.9% Conservative: 33
Best Local Similarity: 71.7% Mismatches: 38
Query Match: 17.3% Indels: 0
DB: 1 Gaps: 0
US-10-681-086-1 (1-3156) x AAP81191 (1-337)
QY 2343 TCGAAGCGCGCATGCTGATATCGGCATATTTGTATCGTCAAGCGCAGAGTCCG 2402
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 82 SerThrAlaProIleGluLysTyrProPheIleThrLysGluGluIleLeuAlaGlyAla 101

QY 2403 AAGCGCGCGCAGCATCTGAATATCGGCACATATTTGTATCGTCAAGCGCAGAGTCCG 2462
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 102 LysArgAlaPheGluAsnLysIleGlyThrTyrCysIleValAlaSerGlyArgGlyPro 121

QY 2463 TCTAACAGAGAAGTGGATCAGGTCGCTAGATCGGTTTCAGGAATTTAAAGAGAGCTATGGA 2522
   ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 122 ThrArgLysAspValAsnValSerGluAlaValGluIleLysAlaLysTyrGly 141

QY 2523 CTGAAGATTGTGCATGCTTGGACTGTTGAACCCAGAGCGCGGCTCAAGAT 2582
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 142 LeuLysValCysAlaCysLeuGlyLeuLeuLysGluGlnAlaGlnLeuLysGlu 161

QY 2583 GCAGGAGTAGACCGCTATAATCAATTTCAATCGTCAAGAGAAACCATTTCAAAATC 2642
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 162 AlaGlyValAspArgTyrAsnHisAsnLeuAsnThrSerGluArgHisHisSerTyrIle 181

QY 2643 ACAACCTCATACATACGATGACAGATCAATACCGTTGAAATTCGAAAGAAATCGGG 2702
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Db 182 ThrThrThrGlnThrTyrGluAspArgValAenThrValGluValValLysLysHisGly 201  
 Qy 2703 CTGTCCTCGCTTTCAGGCGCCATTTATCGGATGAGGAGACGACAGGATGTCATTCAC 2762  
 Db 202 IleSerProCysSerGlyAlaIleGlyMetLysGluThrLysMetAspValValGlu 221  
 Qy 2763 ATCCGCCAAAAGCTTGAAGGCTCTTCAGCGGATTCATTCCTGTGAATTTTTCGATGCA 2822  
 Db 222 IleAlaArgAlaLeuHisGlnLeuAspAlaAspSerIleProValAenPheLeuHisAla 241  
 Qy 2823 ATTGATGCGACCGCTTAGAAGCGCTCAACGAAATTAACCCGCTGTATTGTTTAAAGTG 2882  
 Db 242 IleAspGlyThrLysLeuGluGlyThrGlnAspLeuAenProArgTyrCysLeuLysVal 261  
 Qy 2883 CTGCGCTGTTCCGCTTTATCAATCCATCAAAAGAAATTCGCAATTCGCGAGGAGAGAG 2942  
 Db 262 LeuAlaLeuPheArgTyrMetAsnProSerLysGluIleArgIleSerGlyArgGlu 281  
 Qy 2943 GTCAATCTCCGCACATTCGACGCCATTAGGCTTTACGCCGCAAACTCCATTTTGTGGGA 3002  
 Db 282 ValAenLeuGlyPheLeuGlnProPheGlyLeuTyrAlaAlaAenSerIlePheValGly 301  
 Qy 3003 GACTACTTAAACACTGCGCGGCAAGAGAGAGAGAGATCAATAAATGCTGAGTGATTA 3062  
 Db 302 AspTyrLeuThrThrGluGlyGlnGluAlaAenSerAspTyrArgMetLeuGluAspLeu 321  
 Qy 3063 GGCTTTCGAGTTGATCAGTCGAGAAATGAAG 3095  
 Db 322 GlyPheGluIleGluLeuThrGlnLysGlnGlu 332

RESULT 11

AAW30520  
 ID AAW30520 standard; protein; 387 AA.  
 AC AAW30520;  
 XX 17-OCT-2003 (revised)  
 DT 26-OCT-1998 (first entry)  
 XX Kurthia sp. KAPA synthesetase.  
 DE Kurthia sp. KAPA synthesetase.  
 XX Biotin; vitamin; bioF gene; KAPA synthesetase.  
 XX Kurthia sp. 538-KA26; (DSM 10609).  
 XX EP853127-A2.  
 XX 15-JUL-1998.  
 XX 18-SEP-1997; 97EP-00116237.  
 XX 27-SEP-1996; 96EP-00115540.  
 XX (HOFF ) HOFFMANN LA ROCHE & CO AG F.  
 XX Furuichi Y, Hoshino T, Kimura H, Kiyasu T, Nagahashi Y;  
 WPI; 1998-364652/32.  
 DR N-PSDB; AAV42056.  
 XX New DNA and vectors encoding polypeptides - used for recombinant  
 PT production of biotin.  
 XX Claim 1; Page 20-21; 45pp; English.

This is the KAPA synthesetase encoded by a newly isolated bioF gene (see  
 AAV42056) of Kurthia sp. 538-KA26 (DSM 10609). This enzyme catalyses the  
 conversion of pimelyl CoA to KAPA in the biosynthesis of biotin.  
 CC Polypeptides (see AAW30518-25) encoded by novel bioD, bioA, bioF, bioB,  
 CC bioH, bioI, bioHII and bioC genes (see AAV42054-61) of Kurthia sp. are  
 CC provided, as well as vectors comprising one or more of the genes,  
 CC transformed cells, and a process for the production of biotin that  
 CC comprises cultivating the transformed cells and isolating biotin from the

CC culture medium. Biotin is used in the preparation of pharmaceutical, food  
 CC or feed compositions. The synthesis is an improvement on prior art  
 CC methods involving fermentation of e.g. E. coli and Bacillus sphaericus,  
 CC which have low productivity due to accumulation of DFB, a biotin  
 CC precursor. (Updated on 17-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 387 AA;  
 Alignment Scores:  
 Pred. No.: 5.12e-90 Length: 387  
 Score: 942.00 Matches: 184  
 Percent Similarity: 73.6% Conservative: 81  
 Best Local Similarity: 51.1% Mismatches: 87  
 Query Match: 17.1% Indels: 8  
 Gaps: 4  
 US-10-681-086-1 (1-3156) x AAW30520 (1-387)  
 Qy 37 TGGTTAAACGAGCGGTTAGACAGAATCAAGAAAGCCGCGGTACATCGTAACCTCGCGTCA 96  
 Db 3 TrpGluLysGlu---LeuGluLysIleLysGluGlyGlyLeuTyrArgGlnLeuGlnThr 21  
 Qy 97 ATGATCGAGCGCGGTTCCAGAGAGGAAT-----ATTGATGGCGAAATCAACG 147  
 Db 22 ValGlu-----ThrMetSerAspGlnGlyTyrAlaMetValAenGlyLysLysMetMet 39  
 Qy 148 GTCTGTGCTCAACAACATTTAGGGCTCGCAGCGATAGACGCTTTCATCGATGCGAGCC 207  
 Db 40 MetPheAlaSerAenAenTyrLeuGlyIleAlaAenAspGlnArgLeuIleGluAlaSer 59  
 Qy 208 CAAACAGCATTGACGCAATTTGGGACAGAGCGGTTCCAGCGGTTTACCTTTAAACACAGGCAAT 267  
 Db 60 ValGlnAlaThrGlnArgPheGlyThrGlySerThrGlySerArgLeuThrThrGlyAsn 79  
 Qy 268 TCGGTCTGGCATGAAAGCTAGAAAAGAGATTGCCAGCTTTAAACTGACAGAGCGGCC 327  
 Db 80 ThrIleValHisGluLysLeuGluLysArgLeuAlaGluPheLysGlnThrAspAlaAla 99  
 Qy 328 CTGCTGTTTTCGAGCGGTTACTTGGCCAATGTCGGTCTCCTTTCATCTCCGCCAGAAAG 387  
 Db 100 IleValLeuAenThrGlyTyrMetAlaAsnIleAlaAlaLeuThrThrLeuValGlySer 119  
 Qy 388 GAAGATGTCATTTTAAGTGACCGCTCAATCATGCAAGTATGATCGAGCGGCTGCCGACTT 447  
 Db 120 AspAspLeuLeuLeuSerAspGluMetAsnHisAlaSerIleIleAspGlyCysArgLeu 139  
 Qy 448 TCTAAGGCTGATCAGTTGTTTATCGGCATATTGATGATGATGATGATGATGATGATGATG 507  
 Db 140 SerArgAlaGluThrIleIleTyrArgHisAlaAspLeuLeuAspLeuLeuMetLysLeu 159  
 Qy 508 AATCAACACAGCGTTATCAGCGCGGTTTATCGTAACAGCGGAGTATTCAGCATGGAT 567  
 Db 160 GlnIleAenThrArgTyrArgLysArgIleIleValThrAspGlyValPheSerMetAsp 179  
 Qy 568 GGCACAAATCGCCCTCTTGATCAGATCATCTCATCTCGCAACCGCTATCATGCTCTCGTG 627  
 Db 180 GlyAspIleAlaProLeuProGlyIleValGluLeuAlaLysArgTyrAspAlaLeuVal 199  
 Qy 628 GTCGTTGATGATGCCACGCAAGAGGATTTGGCGGATTCGGGACAGGAGGAGGAGTAA 687  
 Db 200 MetValAspAspAlaHisAlaThrGlyValLeuGlyLysAspGlyArgGlyThrSerGlu 219  
 Qy 688 TACTTTCGTTGTTGTCCTCC-----GACATTGTTATCGGCACCTTAAAGCAAGCTGTTGCG 741  
 Db 220 HisPheGlyLeuLysGlyLysIleAspIleGluMetGlyThrLeuSerLysAlaValGly 239  
 Qy 742 GCGGAAGAGGATTTTGGCGGAGGATCAGCGGCTCTTCATCGCACTTTTTCGTAACCATGCC 801  
 Db 240 AlaGluGlyGlyTyrIleAlaGlySerArgSerLeuValAspTyrValLeuAsnArgAla 259  
 Qy 802 AGAACATTATCTTTAAACCGCTATTTCCGCCAGCCAGCTGTGCGGCTGCTCAGAGGCT 861  
 Db 260 ArgPropheValPheSerThrAlaLeuSerAlaGlyValValAlaAlaSerAlaLeuThrAla 279



Db 282 ArgGluGlyIleSerIleIleGlnAenGluProGluAArgLysGlnLeuLeuLysAen 301  
Qy 913 ATCAGCATGATCAGAACAGCTGTAAGAATATGGTTATGCTGAAGGAGATCACACA 972  
Db 302 AlaGlnTyrLeuArgLeuLysLeuGluSerGlyPheValMetLysGluGlyGluThr 321  
Qy 973 CCGATTATTCCTGTAGTCATTGGCGATGCCCATAAACGGTCCTATTTGCTGAAACCTG 1032  
Db 322 ProfileSerLeuIleIleGlySerHisGluAlaMetGlnPheSerAlaLysLeu 341  
Qy 1033 CAGGCAAGGAATTTATGCTCTCCCATTCGGCCCAACCGTTGCGCCGGTGAAGC 1092  
Db 342 LeuAepGluGlyValPheIleProAlaIleArgProThrValProLysGlySerSer 361  
Qy 1093 CGGATTCGA 1101  
Db 362 ArgLeuArg 364

RESULT 13  
AAW30521  
ID AAW30521 standard; protein; 338 AA.

XX AC  
XX AAW30521;

XX 17-OCT-2003 (revised)  
XX 26-OCT-1998 (first entry)

XX Kurthia sp. DTB synthase.

XX DE Biotin; vitamin; bioB gene; DTB synthase.

XX KW Kurthia sp. 538-KA26; (DSM 10609).

XX OS EP853127-A2.

XX PN 15-JUL-1998.

XX PP 18-SEP-1997; 97EP-00116237.

XX PR 27-SEP-1996; 96EP-00115540.

XX PA (HOFF ) HOFFMANN LA ROCHE & CO AG F.

XX PI Furuichi Y, Hoshino T, Kimura H, Kiyasu T, Nagahashi Y;

XX DR WPI; 1998-364652/32.

XX PT DR N-PSDB; AAW42057.

XX PT New DNA and vectors encoding polypeptides - used for recombinant  
production of biotin.

XX PS Claim 1; Page 22-23; 45pp; English.

XX This is the DTB synthase encoded by a newly isolated bioB gene (see  
CC AAV42057) of Kurthia sp. 538-KA26 (DSM 10609). This enzyme catalyses the  
CC conversion of DTB to biotin. Polypeptides (see AAW30518-25) encoded by  
CC novel bioB, bioA, bioF, bioB, bioH, bioI, bioJ, and bioC genes (see  
CC AAV42054-61) of Kurthia sp. are provided, as well as vectors comprising  
CC one or more of the genes, transformed cells, and a process for the  
CC production of biotin that comprises cultivating the transformed cells and  
CC isolating biotin from the culture medium. Biotin is used in the  
CC preparation of pharmaceutical, food or feed compositions. The synthesis  
CC is an improvement on prior art methods involving fermentation of e.g.  
CC E.coli and Bacillus sphaericus, which have low productivity due to  
CC accumulation of DTB, a biotin precursor. (Updated on 17-OCT-2003 to  
CC standardise OS field)

XX SQ Sequence 338 AA;

Alignment Scores:

Pred. No.: 2.2e-80 Length: 338  
Score: 851.00 Matches: 160  
Percent Similarity: 74.9% Conservative: 46

Best Local Similarity: 58.2% Mismatches: 51  
Query Match: 15.4% Indels: 18  
DB: 2 Gaps: 1  
US-10-681-086-1 (1-3156) x AAW30521 (1-338)  
Qy 2307 TATAGTGGATGTGTCAAAACGCATACCATTTTGAATTCGAAA----- 2348  
Db 60 TyrGlyLysLysValLysLeuAenMetIleIleAenThrLysSerGlyLeuCysProGlu 79  
Qy 2349 -----GCGCCGATTGACTTTCACCGATG 2372  
Db 80 AspCysGlyTyrCysSerGlnSerIleValSerGluAlaProIleAspLysTyrAlaTrp 99  
Qy 2373 GTGAATTAAGAAACGCTGCTTGAAGCGCGAAGCGCGACGATCTCAATATCGGCACA 2432  
Db 100 LeuThrLysGluLysIleValGluGlyAlaGlnGluSerIleArgArgLysAlaGlyThr 119  
Qy 2433 TATTGTATCGTGGCAAGCGCGACAGAGTCCGTCTAACAGAGAAGTGATCAGTCTAGAT 2492  
Db 120 TyrCysIleValAlaSerGlyArgProThrAenArgGluIleAenPheIleGlu 139  
Qy 2493 GCGTTTCAGAAATTAAGAGACGTATGACCTGAAGATTGTCATGCTTGGATGTTG 2552  
Db 140 AlaValLysGluIleArgGluThrThrAspLeuLysIleCysCysLeuGlyPheLeu 159  
Qy 2553 AAGCCAGACGAGCGAGCGGCTCAAAGATCCAGAGTAGACCGCTATAATCATATTG 2612  
Db 160 AenGluThrHisAlaSerLysLeuAlaGluAlaGlyValHisArgTyrLysHisAenLeu 179  
Qy 2613 AATACGTCACAGAGAAACCATTCACCACTCACATACATACATACATACATACATG 2672  
Db 180 AenThrSerGlnAspAenTyrLysAenIleThrSerThrHisThrTyrGluAspArgVal 199  
Qy 2673 AATACGTTGAAATCGCAAAAGAAATCGGGGCTGTCTCGTGTTCAGCGCCCATTCG 2732  
Db 200 AspThrValGluAlaValLysGluAlaGlyMetSerProCysSerGlyAlaIlePheGly 219  
Qy 2733 ATGAAGGAGACGAAACAGGATGTCATTGACATCCGCAAAAGCTTGAAGGCTTTGAC 2792  
Db 220 MetAenGluSerAenGluAlaValGluIleAlaLeuSerLeuArgSerLeuAspAla 239  
Qy 2793 GATTCCATTCTGTGAATTTTTCATGCAATTCATGTCAGCGCGCTTAGAAGCGCTCAAC 2852  
Db 240 AspSerIleProCysAenPheLeuAenAlaIleAspGlyThrProLeuGluGlyThrSer 259  
Qy 2853 GAATTAACCCGCTGTATTGTTTAAAGTGTGCGCTGTTCCTGTTTATCAATCCATCA 2912  
Db 260 GluLeuThrProThrLysCysLeuLysLeuIleSerMetMetArgPheValAenProSer 279  
Qy 2913 AAAGAAATTCGATTTCCGGAGGAGAGAGATCAATCTCCGCACATTCAGCCATTAGGG 2972  
Db 280 LysGluIleArgLeuAlaGlyArgGluValAenLeuArgSerMetGlnProMetAla 299  
Qy 2973 CTTTACGCGCAAACTCCATTTTGTGCGAGACTACTTAACAACCTCCGCGCAAGAGAG 3032  
Db 300 LeuTyrAlaAlaAenSerIlePheValGlyAspTyrLeuThrThrAlaGlyGlnGluPro 319  
Qy 3033 ACGGAGGATCATAAATCTGAGTGATTTAGGCTTTGAAGTTGAA 3077  
Db 320 ThrAlaAepTyrGlyIleIleGluAspLeuGlyPheGluIleGlu 334

RESULT 14

ADF41704

ID ADF41704 standard; protein; 392 AA.

XX AC ADF41704;

XX DT 12-FEB-2004 (first entry)

XX DE Bacillus subtilis KdL amino acid sequence SEQ ID NO:26.

XX KW Bacillus; sbo; slr; ybcO; csn; spollisa; sigB; phrC; rapA; CasS; trpA;





RESULT 15  
ADS44705  
ID ADS44705 standard; protein; 392 AA.  
XX  
AC ADS44705;  
XX  
XX 02-DEC-2004 (first entry)  
XX  
XX Bacterial polypeptide #23135.  
XX  
XX Recombinant DNA construct; transformed plant; improved plant property;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;  
KW pathogen tolerance; pest tolerance; plant disease resistance;  
KW cell cycle pathway modification; plant growth regulator;  
KW homologous recombination; seed oil yield; protein yield; carbohydrate;  
KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;  
KW bacterial polypeptide.  
XX  
XX Bacteria.  
XX  
XX US2003233675-A1.  
XX  
XX 18-DEC-2003.  
XX  
XX 20-FEB-2003; 2003US-00369493.  
XX  
XX 21-FEB-2002; 2002US-0360039P.  
XX  
XX (CAOY/) CAO Y.  
XX (HINK/) HINKLE G J.  
XX (SLAT/) SLATER S C.  
XX (CHEN/) CHEN X.  
XX (GOLD/) GOLDMAN B S.  
XX  
XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;  
XX WPI; 2004-061375/06.  
XX  
XX New recombinant DNA construct comprising a promoter positioned to provide  
XX for expression of a polynucleotide encoding a polypeptide from a  
XX microbial source, useful for producing plants with improved properties.  
XX  
XX Claim 1; SEQ ID NO 23135; 122pp; English.  
XX  
XX The invention relates to a recombinant DNA construct comprising a  
XX promoter functional in a plant cell, where the promoter is positioned to  
XX provide for expression of a polynucleotide encoding a polypeptide from a  
XX microbial source. The invention also relates to a transformed plant  
XX comprising the recombinant DNA construct and a method of producing a  
XX transformed plant having an improved property. The plant is a crop plant  
XX such as maize or soybean. The method of producing a transformed plant  
XX having an improved property comprises transforming a plant with the  
XX recombinant DNA construct and growing the transformed plant, where the  
XX polynucleotide or polypeptide is useful for improving plant properties.  
XX The recombinant DNA construct is useful for producing plants with  
XX improved plant properties, e.g. improved cold, heat or drought tolerance,  
XX tolerance to herbicides, extreme osmotic conditions, pathogens or pests,  
XX increased resistance to plant disease, better growth rate by modification  
XX of the cell cycle pathway with plant growth regulators, increased rate of  
XX homologous recombination, modified seed oil or protein yield and/or  
XX content, improved yield by modification of carbohydrate, nitrogen or  
XX phosphorus use and/or uptake, by modification of photosynthesis or by  
XX providing improved plant growth and development under at least one stress  
XX condition, improved lignin production or improved galactomannan  
XX production. This sequence represents a bacterial polypeptide used in the  
XX scope of the invention. Note: The sequence data for this patent did not  
XX form part of the printed specification but was obtained in electronic  
XX format from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).  
XX  
XX Sequence 392 AA;

Alignment Scores:

Pred. No.: 1.71e-79 Length: 392

```

Db      327 ILeGlyAepGluGlyValAlaLysGlnPheSerAepGlnLeuLeuSerArgGlyValPhe 346
Qy      1051 GCTCCTGCCATTCCGCCCGCCCAACCGTTCCGCCCGGTGAAGCCGGATTCCGAAGCTTGGGC 1110
Db      347 AlaGlnSerIleValPheProThrValAlaLysGlyLysAlaArgIleArgThrIleIle 366
Qy      1111 AGCAGGTCGAGATCAGGGAATGAG 1134
Db      367 ThrAlaGluHisThrLysAepGlu 374

```

Search completed: February 10, 2006, 02:33:17  
 Job time : 485.5 secs

*This Page Blank (uspto)*

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: February 10, 2006, 02:33:31 ; Search time 21.4 Seconds

(without alignments)  
2837.945 Million cell updates/sec

Title: US-10-681-086-1

Perfect score: 5517

Sequence: 1 ggatccacagggttacgacg.....atcggtatgatgtcgaaattc 3156

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 566832

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlbp  
-Q=/abses/ABSSWEB\_spool/US10681086/runat\_09022006\_140517\_11758/app.query.fasta\_1  
-DB=PIR -QFMT=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abs02p  
-USER=US10681086 @CGN\_1\_1\_63 @runat\_09022006\_140517\_11758 -NCPU=6 -ICPU=3  
-NO MMAP -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: PIR:80:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1836	33.3	389	2 F69594	8-amino-7-oxononan
2	1350	24.5	256	2 S09565	kanamycin nucleoti
3	1327	24.1	253	2 B24456	kanamycin nucleoti
4	1317	23.9	335	2 D69594	biotin synthase (E
5	1103	20.0	333	2 D83868	biotin synthase bi
6	971	17.6	332	1 J50274	biotin synthase (E
7	929	16.8	389	1 JQ0512	8-amino-7-oxononan
8	843	15.3	392	2 G69647	glycine C-acetyltr
9	793	14.4	335	2 B90044	biotin synthase [i
10	712	12.9	395	2 C84138	8-amino-7-oxononan
11	698.5	12.7	372	2 A64462	8-amino-7-oxononan
12	692	12.5	398	2 F75017	probable glycine C
13	688	12.5	386	2 AE1853	8-amino-7-oxononan
14	687	12.5	383	2 H89822	hypothetical prote

#### ALIGNMENTS

##### RESULT 1

F69594

8-amino-7-oxononanoate synthase (EC 2.3.1.47) bioF - Bacillus subtilis

C;Species: Bacillus subtilis

C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004

C;Accession: F69594

R;Kunze, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997

A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, Iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F. Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle Rieger, M.; Rivolto, C.; Rocha, B.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Seginchi, J.; Sekowaka, A.; Seror, akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yamamoto, K.; Yata, K.; Yoshida, K. A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A. A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis. A;Reference number: A69580; MUID:98044033; PMID:9384377

A;Accession: F69594

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-389 <KUN>

A;Cross-references: UNIPROT:P53556; UNIPARC:UPI0000060924; GB:Z99119; GB:AL009126; NID:G

A;Experimental source: strain 168

C;Genetics:

A;Gene: bioF

C;Superfamily: 5-aminolevulinatase synthase; glycine C-acetyltransferase homology

C;Keywords: acyltransferase; coenzyme A; phosphoprotein; pyridoxal phosphate

F;44-377/Domain: glycine C-acetyltransferase homology <GCA>

F;237/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Alignment Scores:

Pred. No.: 9.13e-133

Score: 1836.00

Length: 389

Matches: 359

Percent Similarity: 100.0% Conservative: 1  
Best Local Similarity: 99.7% Mismatches: 0  
Query Match: 33.3% Indels: 0  
DB: 2 Gaps: 0

US-10-681-086-1 (1-3156) x F69594 (1-389)

Qy 22 TTGAAGATTGATCTCTGCTTAAACGAGCGGTAGACAGAAATCAAGAGCGCGGTACAT 81  
:::|||||  
Db 1 MetLysIleAaspSerTrpLeuAsnGluArgLeuAaspArgMetLysGluAlaGlyValHis 20  
|||  
Qy 82 CGTAACCTCGCGTCAATGATGAGCGCGGTTCAGAGAGAAATATTGATGGCGAAAT 141  
|||  
Db 21 ArgAsnLeuArgSerMetAaspGlyAlaProValProGluArgAsnIleAaspGlyGluAsn 40  
|||  
Qy 142 CAACGGTCTGCTCTCAAAACAAATATTATGAGCTCGCAAGCATAGACGTTTATCAT 201  
|||  
Db 41 GlnThrValTrpSerSerAsnAsnTyrLeuGlyLeuAlaSerAaspArgLeuIleAasp 60  
|||  
Qy 202 GCAGCCCAACAGCATTCAGCAATTTGGGACAGGACGCGGTTCACGTTTAAACGACA 261  
|||  
Db 61 AlaAlaGlnThrAlaLeuGlnPheGlyThrGlySerSerGlySerArgLeuThrThr 80  
|||  
Qy 262 GGCATTCGCGTCTGCGATGAAAGCTAGAAAAGAGATTGCCAGCTTTTAAACTGCAGAA 321  
|||  
Db 81 GlyAsnSerValTrpHisGluLysLeuGluLysLysIleAlaSerPheLysLeuThrGlu 100  
|||  
Qy 322 GCGGCGCTGCTGTTTTTCAGCGGTTACTTGGCCAAATGTCGGTGTCCTTTCATCTTGC 381  
|||  
Db 101 AlaAlaLeuLeuPheSerSerGlyTyrLeuAlaAsnValGlyValLeuSerSerLeuPro 120  
|||  
Qy 382 GAAAGGAGAGATGTCATTTTAAGTGACCGAGCTCAATCATGCAAGTATCATCGAGCGTGC 441  
|||  
Db 121 GluLysGluAaspValIleLeuSerAaspGlnLeuAsnHisAlaSerMetIleAaspGlyCys 140  
|||  
Qy 442 CGACTTTCTAAGCGTGATACAGTGTGTTTATCGGCATATTGATGATGAATGATCTTGA 501  
|||  
Db 141 ArgLeuSerLysAlaAaspThrValValTyrArgHisIleAaspMetAsnAaspLeuAsn 160  
|||  
Qy 502 AAGTGAATGAAACACAGCGTTATCAGCGCGTGTATTCGTAACAGACGGAGTATTCAGC 561  
|||  
Db 161 LysLeuAsnGluThrGlnArgTyrGlnArgArgPheIleValThrAaspGlyValPheSer 180  
|||  
Qy 562 ATGATGACCAATCGCCCTCTTGATCAGATCATCTCAGTTCGGAACGCTATCATGCC 621  
|||  
Db 181 MetAaspGlyThrIleAlaProLeuAaspGlnIleSerLeuAlaLysArgTyrHisAla 200  
|||  
Qy 622 TTCGTGTCGTGATGATGCCACGCAACAGGAGTTTTGGCGGATTCGGGCAACAGGACG 681  
|||  
Db 201 PheValValValAaspAaspAlaHisAlaThrGlyValLeuGlyAaspSerGlyGlnGlyThr 220  
|||  
Qy 682 AGTGAAATCTTTGTTGTTGTCGCCACATTTGTTATCGGCACCTTAAGCAAAAGCTGTTGC 741  
|||  
Db 221 SerGluTyrPheGlyValCysProAaspIleValIleGlyThrLeuSerLysAlaValGly 240  
|||  
Qy 742 GCGGAAGAGGTTTTGGCGGAGGATCAGCGGTCTTCATCGACTTTTCTGCTCAACCATGCC 801  
|||  
Db 241 AlaGluGlyPheAlaAlaGlySerAlaValPheIleAaspPheLeuLeuAsnHisAla 260  
|||  
Qy 802 AGAATTTATCTTTTCAAAACGCTATTCCGCCAGCCAGCTGTGCGGCTGTCTCAGAGGCT 861  
|||  
Db 261 ArgThrPheIlePheGlnThrAlaIleProAlaSerCysAlaAlaHisGluAla 280  
|||  
Qy 862 TTCACATCATTTGAAGCCAGCGGGAAGAAACAGACAGCTTTTATTTCTTATATCGCATG 921  
|||  
Db 281 PheAsnIleIleGluAlaSerArgGluLysArgGlnLeuLeuPheSerTyrIleSerMet 300  
|||  
Qy 922 ATCAGAACCGAGTCTCAAGAAATATGGTTATGTGTTGAAGGAGATCACACACCATATT 981  
|||  
Db 301 IleArgThrSerLeuLysAsnMetCylTyrValValLysGlyAaspHisThrProIleIle 320  
|||  
Qy 982 CCTGTAGTTCATTTGGCGGATGCCCATAAACGGTCTTATTTGCTTGAATAAACTCGAGGCAAG 1041  
|||

Db 321 ProValValIleGlyAaspAlaHisLysThrValLeuPheAlaGluLysLeuGlnGlyLys 340  
Qy 1042 GGAATTTTATGCTCTGTCATTCGCGCAACCGCTTGGCGGGTGAAGCGCGATTCGA 1101  
|||  
Db 341 GlyIleTyrAlaProAlaIleArgProThrValAlaProGlyGluSerArgIleArg 360  
|||

RESULT 2  
S09565  
kanamycin nucleotidyltransferase (EC 2.7.7.-) [imported] - Staphylococcus aureus plasmid  
A:Alternate names: eomycin resistance protein; kanamycin inactivating enzyme; kanamycin  
C:Species: Staphylococcus aureus  
C:Date: 19-Mar-1997 #sequence revision 25-Apr-1997 #text\_change 09-Jul-2004  
C:Accession: S09565; T4127; A24456  
R:Minton, N.P.; Swinfield, T.J.; Brehm, J.K.; Oultram, J.D.  
Nucleic Acids Res. 18, 1651, 1990  
A:Title: The gram-positive cloning vector pBD64 arose by a 1844 bp deletion of pC194 der  
A:Reference number: S09565; MUID:90221915; PMID:2326208  
A:Accession: S09565  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-256 <MIN>  
A:Cross-references: UNIPROT:O87369; UNIPARC:UPI00000D474F; EMBL:X51450; NID:957998; PID:  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1990  
R:Ito, T.; Katayama, Y.; Hiramatsu, K.  
Antimicrob. Agents Chemother. 43, 1449-1458, 1999  
A:Title: Cloning and nucleotide sequence determination of the entire mec DNA of pre-meth  
A:Reference number: 222733; MUID:99278010; PMID:10348769  
A:Accession: T4127  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-256 <ITO>  
A:Cross-references: UNIPARC:UPI00000D474F; EMBL:D86934; PIDN:BAA82229.1  
A:Experimental source: strain N315  
R:Matsumura, M.; Katakura, Y.; Imanaka, T.; Aiba, S.  
J. Bacteriol. 160, 413-420, 1984  
A:Title: Enzymatic and nucleotide sequence studies of a kanamycin-inactivating enzyme en  
A:Reference number: A24456; MUID:85006820; PMID:6090428  
A:Accession: A24456  
A:Molecule type: DNA  
A:Residues: 'M', 5-256 <MAT>  
A:Cross-references: UNIPARC:UPI000000131C  
C:Genetics:  
A:Gene: aadD  
A:Genome: plasmid pBD64  
C:Superfamily: Staphylococcus kanamycin nucleotidyltransferase  
C:Keywords: antibiotic resistance; nucleotidyltransferase

Alignment Scores:  
Pred. No.: 1.68e-95 Length: 256  
Score: 1350.00 Matches: 256  
Percent Similarity: 100.0% Conservativity: 0  
Best Local Similarity: 100.0% Mismatch: 0  
Query Match: 24.5% Indels: 0  
DB: 2 Gaps: 0

US-10-681-086-1 (1-3156) x S09565 (1-256)

Qy 1570 ATGAGATAGTGAATGACCAATTAATGACTAGACAGAAAGATGATGTTTCAT 1629  
Db 1 MetArgIleValAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHis 20  
|||  
Qy 1630 GAAATTAAGCAACAAATATTGGATAATATGCGGATGATGTTAAGGCTATTGGTGTATTAT 1689  
|||  
Db 21 GluIleLysGluArgIleLeuAaspLysTyrGlyAaspValLysAlaIleGlyValTyr 40  
|||  
Qy 1690 GGCTCTCTGTCGTCTCAGACTGATGGCCCTATTTCGGATATTGAGATGATGTGTCTATG 1749  
|||  
Db 41 GlySerLeuGlyArgGlnThrAaspGlyProTyrSerAaspIleGluMetCysValMet 60  
|||  
Qy 1750 TCAACAGAGGACAGAGTTCAGCCATGGAATGGCAACCGGTGAGTGGAGGTGGAAGTGG 1809  
|||  
Db 61 SerThrGluGluAlaGluPheSerHisGluTrpThrThrGlyGluTrpLysValGluVal 80  
|||

[illegible]

### RESULT 3

B24456  
kanamycin nucleotidyltransferase (EC 2.7.7.-) - *Bacillus* sp. plasmid pTB913  
N;Alternate names: kanamycin inactivating enzyme; kanamycin resistance protein  
C;Species: *Bacillus* sp.  
C;Date: 31-Dec-1990 #sequence revision 20-Feb-1995 #text\_change 09-Jul-2004  
C;Accession: B24456; C24456; S05985  
R;Matsumura, M.; Katakura, Y.; Imanaka, T.; Aiba, S.  
J. Bacteriol. 160, 413-420, 1994  
A;Title: Enzymatic and nucleotide sequence studies of a kanamycin-inactivating enzyme en  
A;Reference number: A24456; MUID:85006820; PMID:6090428  
A;Accession: B24456  
A;Molecule type: DNA  
A;Residues: 1-253 <MA1>  
A;Cross-references: UNIPROT:P05058; UNIPARC:UPI000012DC09; GB:X03409; GB:K02552; NID:940  
A;Accession: C24456  
A;Molecule type: protein  
A;Residues: 1-7 <MA2>  
A;Cross-references: UNIPARC:UPI00001790DB  
R;van der Lelie, D.; Bron, S.; Venema, G.; Oskam, L.  
Nucleic Acids Res. 17, 7283-7294, 1989  
A;Title: Similarity of minus origins of replication and flanking open reading frames of  
A;Reference number: S05980; MUID:90016790; PMID:2677995  
A;Accession: S05985  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 'MRIV', 2-253 <VAN>  
A;Cross-references: UNIPARC:UPI000016E92D; EMBL:X15670; NID:g40078; PIDN:CAA33715.1; PID  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1989  
C;Genetics:  
A;Gene: knt  
A;Genome: plasmid  
A;Start codon: GTG  
C;Superfamily: Staphylococcus kanamycin nucleotidyltransferase  
C;Keywords: antibiotic resistance; nucleotidyltransferase  
F;1-253/Product: kanamycin nucleotidyltransferase #status experimental <MAT>

Alignment Scores:			
Pred. No.:	9.78e-94	Length:	253
Score:	1327.00	Matches:	251
Percent Similarity:	99.6%	Conservative:	1
Best Local Similarity:	99.2%	Mismatches:	1
Query Match:	24.1%	Indels:	0
DB:	2	Gaps:	0
US-10-681-086-1 (1-3156) x B24456 (1-253)			
QY	1579	GTGAATGGACCAATAATAATGACTAGAGAAGAAGATGAAGATTGTTTCATGAAATTAAG	1638
DB	1	MetAsnGlyProIleIleMetThrArgGluGluArgMetIysIleValHisGluIleIys	20
QY	1639	GAACGAATATTGGATAAATATGGGATGATGTTAAAGCTATTGGTGTTTATGGCTCTCTT	1698
DB	21	GluArgIleLeuAspIysTyrGlyAspAspValIysAlaIleGlyValTyrGlySerLeu	40
QY	1699	GGTCGTCAGACTGATGGGCCCTATTCCGATATTTCAGATGATGTGTCTCATGTCTCAACAGAG	1758
DB	41	GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu	60
QY	1759	GAACGAGATTACGCCATGAATGCACAACCGGTGAGTGGAAAGGTGGAAAGTGAAATTTTGAT	1818
DB	61	GluAlaGluPheSerHisGluIuTrpThrThrGlyGluIuTrpIysValGluValAsnPheAsp	80
QY	1819	AGCGAAGAGATTCTACTAGATATTATGCATCTCAGGTGGAATCAGATTGCCCGCTTACACAT	1878
DB	81	SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpProLeuThrHis	100
QY	1879	GGTCAATTTTCTCTATTGTCGCGATTTATGATTCAGGTGGATCTTAGAGAAAGTGAT	1938
DB	101	GlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluIysValTyr	120
QY	1939	CAAACTGCTAAATCGGTAGAGCCCAACGTTCCACGATCGGATTTGTGCGCTTATCGTA	1998
DB	121	GlnThrAlaIysSerValGluAlaGlnIlyPheHisAspAlaIleCysAlaLeuIleVal	140
QY	1999	GAAGAGCTGTTTGAATATGCAGGCAATGCGGTAAATATTCGTGTCAGGACCGCAACA	2058
DB	141	GluGluLeuPheGluTyrAlaGlyIysTrpArgAsnIleArgValGlnGlyProThrThr	160
QY	2059	TTTCTACCATCTTGACTGTACAGGTAGCAATGGCAGTGCCATGTTGTGATTTGGTCTGCAT	2118
DB	161	PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis	180
QY	2119	CATCGCATCTGTTATACGACAGCGGCTTCGTCCTTAATCTAAGCAGTATAGCAATCAGAT	2178
DB	181	HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValIysGlnSerAsp	200
QY	2179	CTTCCTTCAGGTTATGACCATCTGTGCGAGTCTGTAATGTCTGTCAACTTTCCGACTCT	2238
DB	201	LeuProSerGlyTyrAspHisIleCysGlnPheValMetSerGlyGlnLeuSerAspSer	220
QY	2239	GAGAAACTTCTGGAATCGCTAGAGATTTTCTGGAATGGGATTCAGGAGTGGACAAACGA	2298
DB	221	GluIysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluIuTrpThrGluArg	240
QY	2299	CACGGATATATAGTGGATGTGTCAAAACGATACCATTT	2337
DB	241	HisGlyTyrIleValAspValSerIysArgIleProPhe	253

## RESULT 4

D69894  
D69894  
biotin synthase (EC 2.8.1.6) bioB [similarity] - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C:Accession: D69594  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo,  
A.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter,  
C.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.;  
Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi

iech, J.; Harwood, C.R.; Henaute, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.; Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, R.; Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowaka, A.; Seron, akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A. A.; Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*. A;Reference number: A69594; MUID:98044033; PMID:9384377  
A;Accession: D69594  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-335 <KUN>  
A;Cross-references: UNIPROT:P53557; UNIPARC:UPI0000060922; GB:Z99119; GB:AL009126; NID:9  
C;Experimental source: strain 168  
C;Genetics:  
A;Gene: bioB  
C;Superfamily: biotin synthetase  
C;Keywords: 2Fe-2S; biotin biosynthesis; iron-sulfur protein; metalloprotein; sulfotrans  
P;65,69,72,201/Binding site: 2Fe-2S cluster (Cys) (covalent) #status predicted

Alignment Scores:  
Pred. No.: 6.13e-93 Length: 335  
Score: 1317.00 Matches: 263  
Percent Similarity: 92.0% Conservatives: 2  
Best Local Similarity: 91.3% Mismatches: 5  
Query Match: 23.9% Indels: 18  
DB: 2 Gaps: 1

US-10-681-086-1 (1-3156) x D69594 (1-335)

QY 2307 TATAGTGGATGTGTCAAACGCGATACCATTTTGAAT----- 2342  
DB 48 TyrGlyLysValLeuAsnMetIleMetAenAlaLysSerGlyLeuCysProGlu 67

QY 2343 -----TCGAAAGCGCGATTTAGCTTACCGGATG 2372  
DB 68 AsnCysGlyTyrCysSerGlnSerAlaIleSerLysAlaProIleGluSerTyrArgMet 87

QY 2373 GTGAATAAGGAACGCTCTTGAAGCGCGAGCGGCGACGATCTGAATATCGGCACA 2432  
DB 88 ValAsnLysGluThrLeuLeuGluGlyAlaLysArgAlaHisAspLeuAsnIleGlyThr 107

QY 2433 TATTGTATCTGGCAGCGGCGAGAGTCCGCTTAACAGAGAGTGGATCGATCGTAGAT 2492  
DB 108 TyrCysIleValAlaSerGlyArgGlyProSerAenArgGluValAspGlnValAlaAsp 127

QY 2493 GCGGTTTACGAGAAATTAAGAGACGCTATGAGATTTGCTTGGACTGTTG 2552  
DB 128 AlaValGlnGluIleLysGluThrTyrGlyLeuLysIleCysAlaCysLeuGlyLeuLeu 147

QY 2553 AAGCAGACGAGCGGAGCGGCTCAAGATCGAGAGTAGCCGCTATAATCATATAATTG 2612  
DB 148 LysProGluGlnAlaLysArgLeuLysAspAlaGlyValAspArgTyrAsnHisAsnLeu 167

QY 2613 AATACGTCACAGAGAACCATTCACATCAACCTCACATACATACGATGACAGAGTC 2672  
DB 168 AsnThrSerGlnArgAsnHisSerAsnIleThrThrSerHisThrTyrAspAspArgVal 187

QY 2673 AATACGTTGAAATCGCAAAAGATTCGGGCTGCTCCGTTGTCAGGCGCCATTTACGG 2732  
DB 188 AsnThrValGluIleAlaLysGluSerGlyLeuSerProCysSerGlyAlaIleIleGly 207

QY 2733 ATGAAGAGAGCAACACAGATGTCTATTCATCGCAAAAGCTTGAAGGCTCTTGAACGC 2792  
DB 208 MetLysGluThrLysGlnAspValIleAspIleAlaLysSerLeuLysAlaLeuAspAla 227

QY 2793 GATTCATCTCTGTAATTTTTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCG 2852  
DB 228 AspSerIleProValAsnPheLeuHisAlaIleAspGlyThrProLeuGluGlyValAsn 247

QY 2853 GAATTAACCGCGCTGTATTGTTAAAGAGTGTGGCGCTGTTCCGTTTTTCAATCCATCA 2912  
DB 248 GluLeuAsnProLeuTyrCysLeuLysValLeuAlaLeuPheArgPheIleAsnProSer 267

QY 2913 AAAGAAATTCGATTTCCGAGGAGAGAGAGTCAATCTCCGACATTCGACCCATTAGGG 2972  
DB 268 LysGluIleArgIleSerGlyGlyArgGluValAsnLeuArgThrLeuGlnProLeuGly 287

QY 2973 CTTTACCGCCCAACTCCATTTTGTCCGAGACTTACTTAACAATCGCGGCGACAGAGGAG 3032  
DB 288 LeuTyrAlaAlaAsnSerIlePheValGlyAspTyrLeuThrThrAlaGlyGlnGluGlu 307

QY 3033 ACGGAGGATCATAAATGCTGAGTGAATTAAGGCTTTGAAGTTGAATCAGTCGAAGAAATG 3092  
DB 308 ThrGluAspHisLysMetLeuSerAspLeuGlyPheGluValGluSerValGluGluMet 327

QY 3093 AAGCTAGTAAAGTCGAAAGC 3116  
DB 328 LysAlaSerLeuSerAlaLysSer 335

RESULT 5  
biotin synthase bioB [imported] - *Bacillus halodurans* (strain C-125)  
C;Species: *Bacillus halodurans*  
C;Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
C;Accession: D83868  
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira  
Nucleic Acids Res. 28, 4317-4331, 2000  
A;Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and  
A;Reference number: A83650; MUID:20512582; PMID:11058132  
A;Accession: D83868  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-333 <STO>  
A;Cross-references: UNIPROT:Q9K26; UNIPARC:UPI000000C3CA9; GB:AP001513; GB:BA000004; NID  
A;Experimental source: strain C-125  
C;Genetics:  
A;Gene: bioB  
C;Superfamily: biotin synthetase

Alignment Scores:  
Pred. No.: 1.64e-76 Length: 333  
Score: 1103.00 Matches: 206  
Percent Similarity: 85.4% Conservatives: 39  
Best Local Similarity: 71.8% Mismatches: 24  
Query Match: 20.0% Indels: 18  
DB: 2 Gaps: 1

US-10-681-086-1 (1-3156) x D83868 (1-333)

QY 2307 TATAGTGGATGTGTCAAACGCGATACCATTTTGAAT----- 2342  
DB 47 TyrGlyLysValLysLeuAsnMetIleMetAsnAlaLysSerGlyPheCysProGlu 66

QY 2343 -----TCGAAAGCGCGATTTAGCTTACCGGATG 2372  
DB 67 AsnCysGlyTyrCysSerGlnSerSerIleSerLysAlaProIleAspAlaTyrProMet 86

QY 2373 GTGAATAAGGAACGCTCTTGAAGCGCGAGCGGCGACGATCTGAATATCGGCACA 2432  
DB 87 ValAsnLysGluThrIleLeuGluGlyAlaLysArgAlaHisGluLeuAsnValGlyThr 106

QY 2433 TATTGTATCTGGCAGCGGCGAGAGTCCGCTTAACAGAGAGTGGATCGATCGTAGAT 2492  
DB 107 TyrCysIleValAlaSerGlyArgGlyProThrAsnArgAspIleAspHisValThrGlu 126

QY 2493 GCGGTTTACGAGAAATTAAGAGACGCTATGAGATTTGCTTGGACTGTTG 2552  
DB 127 AlaValArgGluIleLysAspThrTyrGlyLeuLysIleCysAlaCysLeuGlyLeuLeu 146

QY 2553 AAGCAGACGAGCGGAGCGGCTCAAGATCGAGAGTAGCCGCTATAATCATATAATTG 2612  
DB 147 LysProGluGlnAlaGluGlnLeuLysAlaAlaGlyValAspArgTyrAsnHisAsnVal 166





A;Cross-references: UNIPROT:P22806; UNIPARC:UPI0000126987; GB:M29291; NID:g142592; PIDN:  
A;Experimental source: strain IF03525  
R;Ploux, O.; Marquet, A.  
Biochem. J. 283, 327-331, 1992  
A;Title: The 8-amino-7-oxopelarginate synthase from *Bacillus sphaericus*. Purification an  
A;Reference number: S21284; MUID:92246854; PMID:1575677  
A;Accession: S21284  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-14 <PLO>  
A;Cross-references: UNIPARC:UPI0000172399  
C;Genetics:  
A;Gene: bioF  
C;Function:  
A;Description: catalyzes the condensation of pimelyl-CoA and L-alanine to form 8-amino-7  
C;Superfamily: 5-aminolevulinatase synthase; glycine C-acetyltransferase homology  
C;Keywords: acyltransferase; biotin biosynthesis; coenzyme A; phosphoprotein; pyridoxal  
F;42-377/Domain: glycine C-acetyltransferase homology <GCA>  
F;237/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Alignment Scores:  
Pred. No.: 3.88e-63 Length: 389  
Score: 929.00 Matches: 187  
Percent Similarity: 70.6% Conservative: 67  
Best Local Similarity: 51.9% Mismatches: 100  
Query Match: 16.8% Indels: 6  
DB: 1 Gaps: 2

US-10-681-086-1 (1-3156) x JQ0512 (1-389)

```
Qy 40 TTAACGAGCGGTAGACAGA-----ATGAAGAAGCGCGGTACATCGTAAC 87
Db 1 MetAenAspArgPheArgArgGluLeuGlnValleGluGluGlnGlyLeuThrArgLys 20
Qy 88 CTGCGGTCAATGGATGGAGCGCGGTCCAGAGAGGAAATATTGATGGCGAAATCAACG 147
Db 21 LeuArgLeuPheSerThrGlyAsnGluSerGluValValMetAenGlyLysLysPheLeu 40
Qy 148 GTCTGGTCTCAACAATATTATAGGCTCCAGCGATAGACGTTTGATCGATCGACGC 207
Db 41 LeuPheSerSerAenAsnTyLeuGlyLeuAlaThrAspSerArgLeuLysLysAla 60
Qy 208 CAAACAGCATTTGGACAGGAGGAGCGGTTCACGTTTAAACACAGGCAAT 267
Db 61 ThrGluGlyIleSerLysTyGlyThrGlyAlaGlySerArgLeuThrGlyAsn 80
Qy 268 TCGTCTGGCATGAAAGCTAGAAGAGATTGCCAGCTTTAAACTGACAGAAGCGGCC 327
Db 81 PheAspIleHisGluGlnLeuGluSerGluIleAlaAspPheLysLysThrGluAla 100
Qy 328 CTGCTGTTTCGAGCGGTACTTGGCCAAATGTCGGTGCCTTTCATCTTCCAGAAAG 387
Db 101 IleValPheSerSerGlyTyLeuAlaAsnValGlyValIleSerSerValMetLysAla 120
Qy 388 GAAGATGTCATTTAAGTGACAGCTCAATCATGCAAGTATGTCAGCGGTCCGACTT 447
Db 121 GlyAspThrIlePheSerAspAlaTrpAsnHisAlaSerIleIleAspGlyCysArgLeu 140
Qy 448 TCTAAGCGTCATGACATGTTGTTTATCGGCATATTGATGATGAATGATCTTTGAACAAAGCTG 507
Db 141 SerLysAlaLysThrIleValTyGluHisAlaAspMetValAspLeuGluArgLysLeu 160
Qy 508 AATGAACACAGCGTTATCAGCGCGGTTTATCGTAACAGCGAGTATTTCAGCATGAT 567
Db 161 ArgGlnSerHisGlyAspGlyLeuLysPheIleValThrAspGlyValPheSerMetAsp 180
Qy 568 GGCACAAATCGCCCTCTTGATCAGATCATCTTCACCTTGGCAACCGCTATCGCTTCGTG 627
Db 181 GlyAspIleAlaProLeuProLysIleValGluLeuAlaLysGluTyLysAlaIle 200
Qy 628 GTCGTTGATGATGCCCAACAGGAGTGTGCGCGATTCGCGCAACAGGACAGGATGAA 687
Db 201 MetIleAspAspAlaHisAlaThrGlyValLeuGlyAenAspGlyCysGlyThrAlaAsp 220
```

```
Qy 688 TACTTTGGTGTGTCGCC-----GACATGTGTATCGGCACCTTAAGCAAGCTGTGGC 741
Db 221 TyrPheGlyLeuLysAspGluIleAspPheThrValGlyThrLeuSerLysAlaIleGly 240
Qy 742 GCGAAGAGAGTGTTCGGCAGGATCAGCGGTCTTCATCGACTTTTGTCTCAACCATGCC 801
Db 241 AlaGluGlyPheValSerThrSerIleAlaLysAsnTyLeuLysAsnAla 260
Qy 802 AGAACATTATCTTTCAACCGCTATTCCGCCAGCCAGCTGTGCGGTCTTCACGAGCT 861
Db 261 ArgSerPheIlePheGlnThrAlaLeuSerProSerAlaIleGluAlaAlaAraGluGly 280
Qy 862 TTCACATCATTTAGCCAGCAGGGAACACGACGCTTTATTTCTTATATCAGCATG 921
Db 281 IleSerIleIleGlnAsnGluProGluArgArgLysGlnLeuLysAsnAlaGlnTy 300
Qy 922 ATCAGAACCATGCTGAAGAATATCGGTATGTGTGTAAGGAGATCATCACACCATGATT 981
Db 301 LeuArgLeuLysLeuGluSerGlyPheValMetLysGluGlyGluThrProIleIle 320
Qy 982 CTTGTACTCATTTGCGGATGCCATAAAACCGTCTCTATTGCTGAAAACCTCGAGGCAAG 1041
Db 321 SerLeuIleIleGlyGlySerHisGluAlaMetGlnPheSerAlaLysLeuLeuAspGlu 340
Qy 1042 GGAATTTATGCTCTGCTATTCGCCATTCGCCCGCAACCGTTGCGCGGTGAACCGCATCGA 1101
Db 341 GlyValPheIleProAlaIleArgProProThrValProLysGlySerSerArgLeuArg 360
```

RESULT 8  
G69647  
glycine C-acetyltransferase (EC 2.3.1.29) kbl - *Bacillus subtilis*  
N;Alternate names: 2-amino-3-ketobutyrate CoA ligase kbl  
C;Species: *Bacillus subtilis*  
C;Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C;Accession: G69647  
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter  
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho  
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrati, E.  
Nature 390, 249-256, 1997  
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler  
Iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.  
Koeter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinols  
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel  
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, M.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,  
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowaka, A.; Seror  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpsstra, P.; Tognoni, A.; Tosato, V.; Uchiyama  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.  
A;Authors: Yoshikawa, H.P.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.  
A;Reference number: A69580; MUID:98044033; PMID:9384377  
A;Accession: G69647  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-392 <KUN>  
A;Cross-references: UNIPROT:O31777; UNIPARC:UPI0000060429; GB:Z99112; GB:AL009126; NID:9  
A;Experimental source: strain 168  
C;Genetics:  
A;Gene: kbl  
C;Superfamily: 5-aminolevulinatase synthase; glycine C-acetyltransferase homology  
C;Keywords: acyltransferase; coenzyme A; phosphoprotein; pyridoxal phosphate  
F;44-380/Domain: glycine C-acetyltransferase homology <GCA>  
F;240/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Alignment Scores:  
Pred. No.: 1.56e-56 Length: 392  
Score: 843.00 Matches: 169  
Percent Similarity: 66.8% Conservative: 77  
Best Local Similarity: 45.9% Mismatches: 120  
Query Match: 15.3% Indels: 2  
DB: 2 Gaps: 1

US-10-681-086-1 (1-3156) x G69647 (1-392)

```
QY 37 TGGTTAAACGAGCGGTTAGACAGAATGAAGAAGACGGCGGTACATCGTACCTCGCGGTCA 96
Db 7 PheLeuLysAlaGluLeuAsnSerMetLysGluAsnHisThrTrpGlnAspIleLysGln 26
QY 97 ATGATGAGCGCGGTTCCAGAGAGGAATATTGATGGCGAATAATCAACGGTCTGGTCC 156
Db 27 LeuGluSerMetGlnGlyProSerValThrValAsnHisGlnLysValIleGlnLeuSer 46
QY 157 TCACAACTATTATTAGGCTCGCAGCGATAGAGCTTGTGATCGACGCCCAACAGCA 216
Db 47 SerAsnAsnTrpLeuGlyPheThrSerHisProArgLeuIleAsnAlaGlnGluAla 66
QY 217 TTGCAGCAATTTGGGACAGAAAGCAGCGGTTACGTTTACGACGAGCAATTCGGTCTGG 276
Db 67 ValGlnGlnTrpGlyAlaGlyThrGlySerValArgThrIleAlaGlyThrPheThrMet 86
QY 277 CATGAAAGCTAGAAAAGAGATTGCGAGCTTTAAACTGACAGAAGCGCCCTGCTGTTT 336
Db 87 HisGlnGluLeuGluLysLysLeuAlaAlaPheLysLysThrGluAlaLeuValPhe 106
QY 337 TCAGCGGTTACTTGGCCATGTCGGTGCTCTCTTCATCTCTGCGCAGAAAGAGATGTC 396
Db 107 GlnSerGlyPheThrThrAsnGlnGlyValLeuSerSerIleLeuSerLysGluAspIle 126
QY 397 ATTTTAAGTCAGCAGCTCAATCATGCAAGTATGATCGACGGCTGCCGACTTCTTAAGGCT 456
Db 127 ValIleSerAspGluLeuAsnHisAlaSerIleLeuAspGlyLeuArgLeuThrLysAla 146
QY 457 GATCAGAGTTGTTTATCGGCATATTGATGATGATCTTGAAACAAAGCTGAATGAACA 516
Db 147 AspLysLysValTrpGlnHisValAsnMetSerAspLeuGluArgValLeuArgLysSer 166
QY 517 CAGCGTTATCAGCGCGCTTTTTCGTAAACAGACGGAGTATTACGATCGATGCGCAATC 576
Db 167 MetAsnTrpArgMetArgLeuIleValThrAspGlyValPheSerMetAspGlyAsnIle 186
QY 577 GCCCTCTTCATCAGATCATCTCAGTTGCGAAACGCTATCATCGCTTGGTGGTGTGAT 636
Db 187 AlaProLeuProAspIleValGluLeuAlaGluLysTrpAspAlaPheValMetValAsp 206
QY 637 GATGCCCAACAGAGAGTTTGGCGATTTCGGGACAGGAACGAGTGAATCTTTGGT 696
Db 207 AspAlaHisAlaSerGlyValLeuGlyGluAsnGlyArgGlyThrValAsnHisPheGly 226
QY 697 GTT-----TGTCCGACATTGTTTCGGCACCTTAAGCAAAAGCTGTTGGCGCGGAGGA 750
Db 227 LeuAspGlyArgValHisIleGlnValGlyThrLeuSerLysAlaIleGlyValLeuGly 246
QY 751 GGTTTTGGCGAGATCAGCGGCTTCATCGACTTTTGTGTAACCATGCCAGAACATTT 810
Db 247 GlyTrpAlaAlaGlySerLysValLeuIleAspTrpLeuArgHisLysGlyArgProPhe 266
QY 811 ATCTTTCAACCCGCTATTCGCGCAGCGAGCTGTGCGGCTGCTCAGCAGGCTTTCAACATC 870
Db 267 LeuPheSerThrSerHisProAlaValThrAlaAlaCysMetGluAlaIleAspVal 286
QY 871 ATTGAAGCCAGCGGGAACACAGACAGCTTTTATTTCTTATATCAGCATGATCAGAAC 930
Db 287 LeuLeuGluGluProGluHisMetGluArgLeuTrpGluAsnThrAlaTrpPheLysAla 306
QY 931 AGTCTGAAGAATATGGGTTATGTTGGTGAAGGAGATCACACCGGATTTCTCTGTAGTC 990
Db 307 MetLeuValLysMetGlyLeuThrLeuThrLysSerGluThrProIleLeuProIleLeu 326
QY 991 ATTTGGGATCCCATAAACCGGCTCTATTGCTGAAACAACTGACGGCGCAAGGAATTTAT 1050
Db 327 IleGlyAspGluGlyValAlaLysGlnPheSerAspGlnLeuLeuSerArgGlyValPhe 346
QY 1051 GCTCTGCGCATTCGCGCGGCAACCGTTTGGCGGGTGAAGCGGATTCGAAGCTTGGGC 1110
Db 347 AlaGlnSerIleValPheProThrValAlaLysGlyLysAlaArgIleArgThrIleIle 366
```

```
QY 1111 AGCAGTCGAGATCAGGAATGAG 1134
Db 367 ThrAlaGluHisThrLysAspGlu 374
RESULT 9
B90044
biotin synthase [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: B90044
C:Kuroda, M.; Ohca, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogut
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaico, C.; Sekimizu, K.; I
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: B90044
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-335 <KUR>
A:Cross-references: UNIPROT:Q99RK7; UNIPARC:UPI00000D7667; GB:BA000018; PID:g13702374; P1
A:Experimental source: strain N315
C:Genetics:
A:Gene: bioB
C:Superfamily: biotin synthetase
Alignment Scores:
Pred. No.: 1,03e-52 Length: 335
Score: 793.00 Matches: 149
Percent Similarity: 73.4% Conservative: 50
Best Local Similarity: 55.0% Mismatches: 54
Query Match: 14.4% Indels: 18
DB: 2 Gaps: 2
US-10-681-086-1 (1-3156) x B90044 (1-335)
QY 2319 GTCAAAACCATACACCACTTTTGAATTCGAAGCCCGCATTT----- 2357
Db 48 ValLysLeuAsnMetIleLeuAsnAlaLysSerGlyLysProGluAsnCysGlyTyr 67
QY 2358 -----GAGCTTACCGGATGGTGAATAAGAAACG 2387
Db 68 CysGlyGlnSerArgAspIleLysGlnLysGlnArgTyrAlaLeuIleProGluGln 87
QY 2388 CTGCTTGAAGCGCGAAGCGCGCACGATCTGAATATCGGCACATATTCGTATCGTGCA 2447
Db 88 IleIleAspGlyAlaLysValAlaHisAspAsnHisIleGlyThrTyrCysIleValMet 107
QY 2448 AGCGGACAGAGTCCGCTTAACACAGAGAAGTCGATCAGGTCTGATCGGTTTCAGGAAT 2507
Db 108 SerGlyArgGlyProSerAspLysGluValAspHisIleSerAsnThrValArgThrIle 127
QY 2508 AAAGAACCTAT---GGACTGAAGATTGTGCATGTCTTGGACTCTTGAAGCCAGACGAG 2564
Db 128 LysSerGlnHisProGlnLeuLysIleCysAlaCysLeuGlyLeuThrAsnAspGluGln 147
QY 2565 GCGAAGCGCTCAAGACGAGTAGACCGCTATAATCATATTTGAATAGTACGTCACAG 2624
Db 148 AlaLysLysLeuLysSerAlaGlyValAspArgTyrAsnHisAsnIleAsnThrSerGlu 167
QY 2625 AGAAACCATCAACCTCACATACATACATACATACATACATACATACATACATACAT 2684
Db 168 AsnTyrHisAspAsnValValThrHisSerTyrLysAspArgThrAspThrIleGlu 187
QY 2685 ATCGCAAAAGAAATCGGGGTGTCTCGTGTTCAGGCGCCATTCATCGGATGAAGAGACG 2744
Db 188 LeuMetLysAlaAsnAsnIleSerProCysSerGlyValIleCysGlyMetGlyGluSer 207
QY 2745 AACACAGATGTCATTCACATCGCCAAAGAGCTTGAAGGCTCTTCACCGGATTCATTCCT 2804
Db 208 AsnGlnAspIleValAspMetAlaPheAlaLeuLysGluMetAspAlaAspSerIlePro 227
```

```
Qy 2805 GTGAATTTTTCATGCAATTGATGCGCACGCGGTTAGAGGCGTCAACGAATTAACCCG 2864
Db :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
228 IleenPheLeuHisProIleuGlyThrLysPheGlySerMetAspAspLeuThrPro 247
Qy 2865 CTGTATTGTTTAAAGTCTGCGCGTGTTCGGTGTTCATCAATCCATCAAAAGAAATTCGC 2924
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
248 MetLysCysLeuArgIleValAlaLeuPheArgLeuIleAsnProThrLysGluIleArg 267
Qy 2925 ATTTCCGAGGAGAGAGTCAATCTCCGACATTCGACATTCAGCCATTAGGCGTTTACCGCA 2984
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
268 IleAlaGlyArgGluValAsnLeuArgSerLeuGlnProLeuAlaLeuLysAlaAla 287
Qy 2985 AACTCCATTTTGTCCGAGACTACTTAACAACCTGCCGGCGCAAGAGGACGAGGAGATCAT 3044
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
288 AsnSerIlePheValGlyAspTyrLeuIleThrGlyGlyGlnProAsnGlnLeuAspTyr 307
Qy 3045 AAAATGCTGAGTGATTAGGCTTTGAAGTTGAA 3077
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
308 AspMetIleAsnAspLeuGlyPheGluIleAsp 318

RESULT 10
C84138
8-amino-7-oxononanoate synthase bioF [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Spec: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: C84138
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: C84138
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-395 <STO>
A:Cross-references: UNIPROT:Q9K625; UNIPARC:UPI00000C4332; GB:AP001520; GB:BA000004; NID
A:Experimental source: strain C-125
C:Genetics:
A:Gene: bioF
C:Superfamily: 5-aminolevulinatase synthase; glycine C-acetyltransferase homology
C:Keywords: phosphoprotein; pyridoxal phosphate
F:240/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Alignment Scores:
Pred. No.: 1,78e-46 Length: 395
Score: 712.00 Matches: 153
Percent Similarity: 62.5% Conservative: 74
Best Local Similarity: 42.1% Mismatches: 124
Query Match: 12.9% Indels: 12
DB: 2 Gaps: 5

US-10-681-086-1 (1-3156) x C84138 (1-395)
Qy 37 TGGTTAAAC-----GAGCGGTAGACAGAATGAAAGAGCGCGGTACATCGTAAC 87
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
5 TrpLeuHisAlaIleGluGluLysLeuThrArgLeuLysAspArgGlySerPheArgGln 24
Qy 88 CTGCGGTCAATGGATGGAGCGCGGTTCACAGAGGAATATGTGCGGAAATCAACG 147
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
25 LeuValProThrSerGluAlaLeuProThrLeuThrArgGluAsnCysArgLeuLeu 44
Qy 148 GTCTGGTCTCTCAACAATTTATAGGCTCGCAAGCATAGACGTTTGTATCGATCGTCAGCC 207
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
45 AsnLeuAlaSerAsnAsnTyrLeuGlyIleAlaAspSerLysGluPheIleGluArgThr 64
Qy 208 CAACAGCATTTGCAGCAATTTGGGACAGAGCGGTTCACGTTTAAACGACAGGCAAT 267
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
65 GluGlnLeuAlaSerSerTyrAlaIleGlySerThrAlaSerArgLeuIleIleGlyAsn 84
Qy 268 TCGTCTGGCATGAAGACTAGAAAGAGATTCGCAGCTTTAACTGACAGAGCGGCC 327
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
85 HisProLeuTyrGluGluAlaGluTyrGluLeuThrLysTyrLysLysThrGluAlaAla 104
Qy 328 CTGCTGTTTCGAGCGGTACTTGCCCAATGTCGGTGTCTTCATCTTCCTGCCAGAAAG 387
```

```
Db 105 LeuIlePheGlySerGlyTyrMetAlaAsnValGlyIleIleSerSerIleValGlyArg 124
Qy 388 GAAGATGTCATTTTAAAGTACACAGCTCAATCATGCAAGTATGATGCGGTCGCGACTT 447
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
125 GlyAspAlaValPheSerAspLysLeuAsnHisAlaSerIleValAspGlyCysGlnLeu 144
Qy 448 TCFAAGGCTCATACAGTTGTTTATCCGCATATGATGATGATGATGATGATGATGATGATG 507
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
145 SerArgAlaAspHisLeuArgPheArgHisAsnAspMetAspHisLeuGluThrLeuLeu 164
Qy 508 AATGAACACAGCGCTTATACGCGCGTTCCTTAACACAGCGAGGATTCAGCATGAT 567
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
165 GlnLysSerPro---HisLysGlnLysLeuIleValValAspAlaLeuPheSerMetAsp 183
Qy 568 GGCACAAATCGCCCTCTTGATCAGATCATCTCATCTCGGAAACCGCTATCATGCTTCGTG 627
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
184 GlyAspHisAlaAsnLeuHisAspLeuValThrLeuLysGluArgTyrGlyAlaIleLeu 203
Qy 628 GTCGTTGATGATGCCACGCAACAGGAGTTTTCGGCGATTTCGGCACAAGGACAGTGAA 687
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
204 MetValAspGluAlaHisSerGlyValTyrGlyAlaThrGlyGlyLeuValGlu 223
Qy 688 TACTTTGCTGTGTTGCC-----GACATTTGTTATCGGCACCTTAAGCAAGCTGTTGGC 741
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
224 GluLeuGlyLeuAsnAspArgValAspIleGlnMetGlyThrPheSerLysAlaLeuGly 243
Qy 742 GCGAAGAGAGGTTTTCGGCAGATCAGCGGTCTTCATCGACTTTTTCGTAACCATGCC 801
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
244 SerTyrGlyGlyTyrValAlaGlyAlaLysPheIleGluTyrLeuLeuAsnHisAla 263
Qy 802 AGAACATTTATCTTCAACCGCTATTCGCCAGCCAGCTGTGCGGTGCTCAGGAGCT 861
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
264 ArgSerLeuIlePheThrThrAlaLeuProProTyrIleValAlaSerHisLeuAlaAla 283
Qy 862 TTCAACATCATTTGAAGCCAGC-----AGGAAAAACGACAGCTTTTATTTTCTTAT 912
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
284 LeuGlnIleValGlnGluGlnProTrpArgArgGluLysValGlnValLeuGly----- 301
Qy 913 ATCAGCATGATCAGAACCATGCTGAAGAATATGGTTATGTGTGAAGAGGATCACACA 972
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
302 ---GluArgLeuArgAsnGlyLeuGluGlnLeuGlyPheSerLysGlySerGluSer 320
Qy 973 CCGATTATTCCTGTAGTCATTGGCGATGCCATAAAACGGTCTCTATTTCGTGAAAAACTG 1032
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
321 TyrIleValProValLeuIleGlyAspAsnHisAspLeuLeuValSerGluSerLeu 340
Qy 1033 CAGGCGCAAGGGAATTTATGCTCTCTCCATTCGGCCGCAACCGTTGCGCGGGTGAAGC 1092
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
341 GlnAlaAlaGlyIleAlaAlaIleProValArgProProThrValProArgGlyGluGly 360
Qy 1093 CGGATTTCGA 1101
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
361 ArgIleArg 363

RESULT 11
A64462
8-amino-7-oxononanoate synthase (EC 2.3.1.47) - Methanococcus jannaschii
C:Species: Methanococcus jannaschii
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C:Accession: A64462
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
rsen, J.D.; Sadov, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: A64462
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-372 <BUL>
A:Cross-references: UNIPROT:Q58694; UNIPARC:UPI000012698C; GB:U67570; GB:L77117; NID:928
```

C;Genetics:

A;Map position: FOR1245358-1246476

C;Superfamily: 5-aminolevulinatase synthase; glycine C-acetyltransferase homology

C;Keywords: acyltransferase; coenzyme A; phosphoprotein; pyridoxal phosphate

F;32-363/Domain: glycine C-acetyltransferase homology <GCA>

F;226/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Alignment Scores:

Pred. No.:	1.9e-45	Length:	372
Score:	698.50	Matches:	158
Percent Similarity:	62.6%	Conservative:	66
Best Local Similarity:	44.1%	Mismatches:	113
Query Match:	12.7%	Indels:	21
DB:	2	Gaps:	9

US-10-681-086-1 (1-3156) x A64462 (1-372)

Qy	40	TTAAACGAGCGGTTAGCAGACATGAAGAAGCCGGCTATCATCGTAACCTCGCGTCAATG	99
Db	6	LeuArgargGlulleGluillelleyAsnAnsglyeuYrArGPheLeuArgLysLys	25
Qy	100	GATGGAGCGCGGTTCAGAGAGAAATATTGATCGCGAAAATCAAACGGTGCTGCTCTCA	159
Db	26	Asp-----PspglyValleuAasp-----PheSerSer	34
Qy	160	AACAATTTATTAGGCCTCGCAAGCGATAGACGTTTGATCGATGAGCCCCAACAGCATTTG	219
Db	35	AsnAspTyrlEuCyLSerLysHieProgluValleGluAlaValLysGluGlyLeu	54
Qy	220	CAGCAATTTGGGACAGGAAGCAGCGGTTTCAGTTTAAACGACAGGCAATTCGGTCTGCCAT	279
Db	55	--LysTyrglyAlaGlySerThrGlySerArgLeuthrSerGlyAsn---IleAsnHis	72
Qy	280	GAAGAAGCTAGAAAGAGATGTCAGGCTTTAAACTGACAGAGCGGCCCTGCTGTTTTTCG	339
Db	73	GlnArgLeuGluGluLysIleAlaagLPheLysGluThrGluAArgThrLeuValTy-Ser	92
Qy	340	AGCGGTACTTGGCCAATGTCGGTGTCCTTTTCATCCTTGGCCAGAAAAAGGAGATGTCATT	399
Db	93	SerGlyTyrrAlaThrAsnValGlyValIleSerAlalaLeuCysLySlysGlyAspLeuile	112
Qy	400	TTAAGTCACGAGCTCAATCATGCAAGTATGATCGACGGCTGCCGACTTTCTTAAGCGCTGAT	459
Db	113	LeuSerAspLysLeuAsnHisAlaSerIlelleaspglyCyslyLysLeuSerLysAlaAsp	132
Qy	460	ACAGTTGTTTATCGGCATATTGATATGAATGATCTTGAAAACAAGCTGAATGAACAAC-	516
Db	133	ValLeuiletyrAsnHisCysAspValGluHisLeuThrAsnLeuilleGluGluAsnTrp	152
Qy	517	CAGCGTTATCAGCGCCGTTTTATCGTAAACAGCGAGTATTCAGCATGGATGGCACATC	576
Db	153	GlyLysTyrrAsnAsnLeuPheilleValThrAspGlyValPheSerMetAspGlyAspile	172
Qy	577	GCCCCTCTTGATCAGATCATCTCATTCGCGAAACGCTATCATGCTTCGTGGTCTGGTGGAT	636
Db	173	AlaProlueargAspLeuLysLysIleAlaAspGluPheAsnAlaIleLeuilleleaspp	192
Qy	637	GATGCCACGCAACAGGAGTTTTGGCGATTCGGGACAGGAACGAGTGAATACTTTGGT	696
Db	193	AspAlaHisgLyThrGlyValLeuGlyAasp--GlyArgGlyThrLeuLysHisPheAsn	211
Qy	697	GTTTGTCCCGNAC-----ATTGTTTATCGGCACCTTAGCAAGCTGTGGCGCGGAA	747
Db	212	LeuLysProSerAspAsnIleValGlnIleGlyThrLeuSerLysAlalleGlyGlyLeu	231
Qy	748	GGAGGTTTTGCGGCAAGATCAGCGGCTCTTCATCGACTTTTTCGTGAACCATGCCACAAC	807
Db	232	GlyGlyPheValCysGlylleGluGluValValGluTyrrLeuilleAsnThrSerArgSer	251
Qy	808	TTTATCTTTCAAACCGCTATTTCGCCACGCCAGCTGTGGCGCTGCTACGAGCGTTTCAAC	867
Db	252	PheillePheSerThrAlaLeuProPrOHisValValGluGlyCysIleLysAlaPheGlu	271

Qy	868	ATCATTTGAAGCCAGCAGGAAAACGACGCTTTTATTTCCTATATCATCAGCATGATCAGA	927
Db	272	lleileglulysThrAspIleValLysLeuLysGluN---LysAsnIleLysIleAlaAsn	290
Qy	928	ACCAGTCTGAAGAATATGGGTATGTGGTGAAGAGATCACACACGATTTATTCCTGTA	987
Db	291	LysValPheLysLysTyThrGluPheIleLysGluAspAsnLeuThrProIleTyThrPhe	310
Qy	988	GTCATTGGCGATGCCCATAAACGGTCCTATTGTCCTAAAAAATCTCAGGGCGAAGGAATT	1047
Db	311	IlePheLysGlu-----LysThrMetGluIleAlaGluHisLeuIleLysAsnAsnIle	328
Qy	1048	TATGCTCTCCCATTCGGCGCCCAACCGTTGCGCGGGTGAAACCCGATTCGA	1101
Db	329	PheCysValGlyIleArgTyProThrValProLysGlyLeuGluArgIleArg	346
RESULT 12			
F75017			
probable glycine C-acetyltransferase (EC 2.3.1.29) PAB1244 - Pyrococcus abyssi			
C/Species: Pyrococcus abyssi			
C/Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004			
C/Accession: F75017			
R:anonymous, Genoscope			
submitted to the EMBL Data Library, July 1999			
A/Description: Pyrococcus abyssi genome sequence: insights into archaeal chromo			
A/Reference number: A75001			
A/Accession: F75017			
A/Status: preliminary			
A/Molecule type: DNA			
A/Residues: 1-398 <RAW>			
A/Cross-references: UNIPROT:Q9UY32; UNIPARC:UPT000000345A5; GB:AJ248288; GB:AL09			
A/Experimental source: strain Orsay			
C/Genetics:			
A/Gene: PAB1244			
C/Superfamily: 5-aminolevulinate synthase; glycine C-acetyltransferase homology			
C/Keywords: acyltransferase; coenzyme A; phosphoprotein; pyridoxal phosphate			
F/47-385/Domain: glycine C-acetyltransferase homology <GCH>			
F/244/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted			
Alignment Scores:			
Pred. No.:	6.1e-45	Length:	398
Score:	692.00	Matches:	151
Percent Similarity:	61.8%	Conservative:	74
Best Local Similarity:	41.5%	Mismatches:	133
Query Match:	12.5%	Indels:	6
DB:	2	Gaps:	4

US-10-681-086-1 (1-3156) x F75017 (1-398)

Qy	25	AAGATTGATTCCTCGTTAAACAGCGCGTTAGACAGAATGAAGAAAGCGCGGTACATCGT	84
Db	6	LysLeuAsp---TrpIleLysGluGluLeuGluLeuLysLysGlyLeuTyrVal	24
Qy	85	AACTCGCGTCAATGGATGGAGCGCCGGTTCAGAGAGAGGAATATGTGATGCGCAAAATCAA	144
Db	25	ThrIleArgValLeuGlnSerAlaGlnGlyProTrpValValValAsnGlyLysArgVal	44
Qy	145	ACGTCCTGGTCCTCAAAACAATATTTAGGCTCGCAAGCGCATGACGTTTCATCGATGCA	204
Db	45	LeuAsnMetCysSerAsnAsnTyrLeuGlyLeuAlaAlaHisProLysIleLysGluAla	64
Qy	205	GCCCAACACGATTCGACGAATTTGGACAGGACGACGCGTTCCACGTTTAAACGACACGCG	264
Db	65	AlaIleArgAlaIleLeuAspTyrGlyValGlyAlaGlyAlaValArgThrIleAlaGly	84
Qy	265	AATTTCGCTCGGCATGAAAGACTAGAAAAGAATTCGCCAGCTTTAAACTGACACAGCG	324
Db	85	ThrMetGluLeuHisValGluLeuGluLeuAlaLysPheLysLysArgGluAla	104
Qy	325	GCCCTGCTGTTTCGACGGTTACTTCGGCAATGTCTGGTCTCTTTTCATCTCTCCAGAA	384
Db	105	AlaIleLeuPheGlnSerGlyTyrAsnAlaAsnLeuGlyValIleSerAlaLeuLeuArg	124



**QY** 412 CTCAATCATGCAAGTATGATGCACGGCTGCCAAGCTTTCTAAAGCTGATACAGTTGTATT 471  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
**Db** 121 LeuAsnHisAlaSerIleileAspGlyCyArgLeuSerLysAlaLysIleileArgVal 140  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
**QY** 472 CGGCATATTGATGATGAATGATCTTGAAAACACAGCTGAAT-----CAAACACAGCGT 522  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
**Db** 141 AsnHisSerAspMetCaspLeuArgAlaLysAlaLysGluAlaValGluseryGlyIn 160  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
**QY** 523 TATCAGCGCGCTTTTATCGTAACACAGCGAGTATTACAGCATCGATGCGCAATTCGCCCT 582  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
**Db** 161 TyrAsnLysValMetTyrIleThrAspGlyValPheSerMetCaspGlyAspValAlaLys 180  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
**QY** 583 CTTGATCAGATCATCTCCGAAACGCTATCATGCTTCGTGGTCTCTGATGATGCC 642  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
**Db** 181 LeuProGluIleValGluIleAlaGluGluPheGlyLeuLeuThrTyrValAspAspAla 200  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
**QY** 643 CACGCAACAGGAGTTTGGCGCATTCGGGACAAGCAAGTGAATACATTCTTGGTGTTCGT 702  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
**Db** 201 HisGlySerGlyValMetGly---LysGlyAlaGlyThrValLysHisPheGlyLeuGln 219  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
**QY** 703 CCC-----GACATTGTTATCGGCACCTTAAGCAAGCTGTGGCGCGAAGCAGGTTTT 756  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
**Db** 220 AspLysIleAspPheGlnIleGlyThrLeuSerLysAlaIleGlyValValGlyGlyTyr 239  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
**QY** 757 GCGCAGGATGACGGCTTCATCGACTTTTCTGCTGAACCATGCCAGAACATTATTCCTTT 816  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
**Db** 240 ValAlaGlyThrLysgluLeuileAspTrpLeuLysAlaGlnSerArgPropHeuPhe 259  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
**QY** 817 CAACCGCTATTCGCGCACCGAGCTGTGGCGTCTCAGAGGCTTTCACATCATTTGAA 876  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
**Db** 260 SerThrSerLeuAlaProGlyAspThrLysAlaIleThrGluAlaValLysLysLeuMet 279  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
**QY** 877 GCCAGCAGGAAAAACGACAGCTTTATTTCTTATATCAGCATGATCAGAACCAGCTCG 936  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
**Db** 280 AspSerThrGluLeuHisAspLysLeuTrpAspAsnAlaGlnTyrLeuLysAsnGlyLeu 299  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
**QY** 937 AAGAATATGGTGTATGTGTGTAAGGAGATCATCACACCGATTATTCCTGTAGTCATTCGC 996  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
**Db** 300 SerLysLeuGlyTyrAspThrGlyGluSerGluThrProIleThrProValIleIleGly 319  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
**QY** 997 GATGCCCATMAAACGGTCTCTTCTGTAATACTGCAGGCAAGGGAATTTATGCTCTCT 1056  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
**Db** 320 AspGluLysThrThrGlnGluPheSerLysArgLeuLysAspGluGlyValTyrValLys 339  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
**QY** 1057 GCATTTCGSCGCCCAACGCTTCGCGCGGTGAAGCCGATTCGAAGCTTG 1107  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
**Db** 340 SerIleValPheProThrValProArgGlyThrGlyArgValArgAsnMet 356  
| | | | | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
  
**RESULT 15**  
E71454  
probable glycine C-acetyltransferase (EC 2.3.1.29) PH0292 - Pyrococcus horikoshii  
C:Species: Pyrococcus horikoshii  
C>Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 16-Aug-2004  
C:Accession: E71454  
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.  
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.;  
DNA Res. 5, 55-76, 1998  
A:Title: Complete sequence and gene organization of the genome of a hyper-thermo-  
philic archaeon, Pyrococcus horikoshii strain DsmZ  
A:Reference number: A71000; PMID:98344137; PMID:9679194  
A:Accession: E71454  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-398 <RAW>  
A:Cross-references: UNIPROT:O58030; UNIPARC:UPI00000668AA; GB:AP0000001; NID:9322  
A:Experimental source: strain O73  
A>Note: This accession replaces an interim accession for a sequence replaced by  
C:Genetics:  
A:Gene: PH0292  
C:Keywords: acyltransferase; coenzyme A; phosphoprotein; pyridoxal phosphate  
F:47-385/Domain: Glycine C-acetyltransferase homology <GCA>  
F:244/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted



## Alignment Scores:

Pred. No.: 1.76e-44 Length: 398  
Score: 686.00 Matches: 150  
Percent Similarity: 61.8% Conservative: 75  
Best Local Similarity: 41.2% Mismatches: 133  
Query Match: 12.4% Indels: 6  
DB: 2 Gaps: 4

US-10-681-086-1 (1-3156) x E/1454 (1-398)

```
QY 25 AAGATTGATCTCTGTTAAAGCGGTTAGACAGATGAAGAAGCGCGGTACATCGT 84
Db 6 LysLeuAseP---TrpIleGluGluLeuGluLeuLysLysLysLeuTyrVal 24
QY 85 AACCTGCGGTCAATGATGGAGCGCGGTTCAGAGAGGAATATTGATGGCGAAATCAA 144
Db 25 ThrIleArgValLeuGlnSerAlaGlnGlyProTrpIleValValAsnGlyLysArgVal 44
QY 145 ACGGTCTGCTCCTCAAAACAATTTATAGGCTCGCAAGCGATAGACGTTTGTATCGATCA 204
Db 45 LeuAsnMetCysSerAsnAsnTyrLeuGlyLeuAlaAlaHisProLysIleLysGluAla 64
QY 205 GCCCAACAGCATTCACCAATTTGGGACAGGAGCAGCGGTTCACGTTTACGACAGGC 264
Db 65 AlaIleArgAlaIleLeuAspTyrGlyValGlyAlaGlyAlaValArgThrIleAlaGly 84
QY 265 AATTGCGGTCTGCGATGAAGCTAGAAAAGAGATTGCCAGCTTTAAACTGACAGAGCG 324
Db 85 ThrMetGluLeuHisValGluLeuGluGlnLysLeuAlaLysPheLysLysArgGluAla 104
QY 325 GCCCTGCTGTTTTCGAGCGGTTACTTTGCCAATGTCGCTGTCCTTTCATCTCTGCCAGAA 384
Db 105 AlaIleLeuPheGlnSerGlyTyrAsnAlaAsnLeuGlyAlaIleSerAlaLeuLeuArg 124
QY 385 AAG-----GAAGATGTCATTTAAGTGACGACGCTCAATCATGCAAGTATGATCGACGC 438
Db 125 LysGlyGluAspGlyValPheLeuSerGluGluLeuAsnHisAlaSerIleIleAspGly 144
QY 439 TGCGGACTTCTAAGGCTGATACAGTGTGTTATCGGCATATTGATGATGAATGATCTTGAA 498
Db 145 MetArgLeuSerGlyAlaProLysValIleTyrLysHisLeuAspValAspAspLeuLys 164
QY 499 AACAGCTGAATGAACAACAGCGTTCATCAGCGCGCTTTTATCGTAACACAGCGAGTATTC 558
Db 165 LysLysLeuGluGluAsnLysAspLysLysLysIleIleValThrAspGlyValPhe 184
QY 559 AGCATGATGACACATCCGCCCTCTTGCATCAGATCATCTCCTTCGAAACGCTATCAT 618
Db 185 SerMetAspGlyAspLeuAlaProLeuProGluIleValGluValAlaGluGlnTyrAsp 204
QY 619 GCCTTCGTGTCGTTGATGCCACGACGAGTTTGGGCGATTTCGGGACAGGA 678
Db 205 AlaIleValTyrValAspAspAlaHisGlyGluGlyValLeuGlySerHisGlyArgGly 224
QY 679 ACGAGTGAATCTTTGTTGTTGTCCC-----GACATTTGTTATCGGCACCTTAAGCAAA 732
Db 225 IleValAspHisPheAsnLeuHisAspLysValAspPheGluMetGlyThrLeuSerLys 244
QY 733 GCTGTTGGCGGAGAGGTTTTCGGCGCAGGATCAGCGGTCTTCATCGACTTTTGTCTG 792
Db 245 AlaPheGlyValIleGlyGlyTyrValAlaGlyProGluGluAlaIleGluTyrLeuLys 264
QY 793 AACCATGCCAGACATTTTCTTCAACCGCTATTCGCCCGCAGCTGTGCGGCTGCT 852
Db 265 GlnArgAlaArgProPheLeuPheSerSerAlaMetAsnProProAspValAlaAlaAla 284
QY 853 CACGAGGCTTTCAACATCATTTAAGCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 912
Db 285 IleAlaAlaValGluIleLeuGlnLysSerAspAspLeuValLysLysLeuTipAspAsn 304
QY 913 ATCAGCATGATCAGACACGCTGTAGAGATATGTTATGTTGTTGTTGTTGTTGTTGTTG 972
Db 305 ThrHisPheGlnLysGlyLeuArgAspLeuGlyTyrAspLeuGlyAsnThrLysHis 324
```

```
QY 973 CCGATTATTCTCTAGTCAATTGGCGATGCCATAAAACGGTCTCTATTGTTGCTGAAAAACTG 1032
Db 325 ProIleThrProValMetLeuTyrAspGluLysLeuAlaGlnGluPheSerArgArgLeu 344
QY 1033 ---CAGGGCAAGGAATTTATGCTCTCCATTCCGCCGCGCCCAACCGTTGCGCGGTCGAA 1089
Db 345 TyrGluGluTyrAsnIlePheAlaGlnAlaIleValTyrProThrValProLeuGlyThr 364
QY 1090 AGCCGGATTGCA 1101
Db 365 AlaArgIleArg 368
```

Search completed: February 10, 2006, 03:00:23  
Job time : 129 secs



## Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension .rup) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

**When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.**

**BEST AVAILABLE COPY**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: February 10, 2006, 02:17:36 ; Search time 139.9 Seconds

(without alignments)  
3183.198 Million cell updates/sec

Title: US-10-681-086-1

Perfect score: 5517

Sequence: 1 ggatccacgaggtacgacg.....atcggtatgatgcgaattc 3156

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Delop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 4332886

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODELS=frame+ n2p.model -DEV=xlp  
-Q=/abses/ABSSWEB\_spool/US10681086/runat\_09022006\_140515\_11713/app\_query.fasta\_1  
-DB=UniProt -QFWT=fastan -SUFFIX=n2p.rup -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=bloum62 -TRANS=human40.cdi -LIST=45  
-DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUT=txt -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abses02p  
-USN=US10681086 @CGN 1.1 466 @runat\_09022006\_140515\_11713 -NCPU=6 -ICPU=3  
-NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV TIMEOUT=120  
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DBLOP=6 -DELEXT=7

Database : UniProt 05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1836	33.3	389	1	BIOF_BACSU
2	1676	30.4	371	2	Q8KZM9_BACSU
3	1350	24.5	256	2	Q7BU13_STAAU
4	1350	24.5	256	2	Q6GKR2_STAAU
5	1350	24.5	256	2	O87369_STAAU
6	1350	24.5	256	2	Q7ABD0_STAAU
7	1333	24.2	253	1	KANU_STAAU
8	1333	24.2	253	2	Q7DU50_STAAU
9	1327	24.1	253	1	KANU_BACSP
10	1318	23.9	253	2	Q57514_BACST
11	1317	23.9	335	1	BIOB_BACSU
12	1308	23.7	335	2	Q8KZM7_BACSU
13	1286.5	23.3	379	2	Q65ML1_BACLD
14	1222	22.1	253	2	Q75WE9_92ZLD
15	1222	22.1	333	2	Q65MK9_BACLD
16	1214	22.0	339	2	Q70JZ1_BACAM

17	1103	20.0	333	2	Q9KC26_BACHD	Q9KC26 bacillus ha
18	1050	19.0	333	2	Q5KZNI_GEOKA	Q5KZNI geobacillus
19	971	17.6	332	1	BIOB_BACSH	P19206 bacillus sp
20	942	17.1	387	2	Q9AJN1_9BACL	Q9AJN1 kurtzia sp.
21	929	16.8	389	1	BIOF_BACSH	P22806 bacillus sp
22	851	15.4	338	2	Q9AJN0_9BACL	Q9AJN0 kurtzia sp.
23	843	15.3	392	1	KBL_BACSU	O31777 bacillus su
24	828	15.0	313	2	Q8COB3_STABP	Q8COB3 staphylococ
25	828	15.0	321	2	Q5HKJ7_STABEQ	Q5HKJ7 staphylococ
26	809	14.7	332	2	Q4MMW5_BACEE	Q4MMW5 bacillus ce
27	809	14.7	332	2	Q6HE51_BACHK	Q6HE51 bacillus th
28	809	14.7	332	2	Q635G7_BACCZ	Q635G7 bacillus an
29	809	14.7	332	2	Q81MB3_BACAN	Q81MB3 bacillus an
30	808	14.6	332	2	Q73112_BACC1	Q73112 bacillus ce
31	804	14.6	332	2	Q818X3_BACCR	Q818X3 bacillus ce
32	804	14.6	391	2	Q65JE6_BACLD	Q65JE6 bacillus li
33	793	14.4	335	2	Q6G6P6_STAAS	Q6G6P6 staphylococ
34	733	14.4	335	2	Q6GE08_STAAU	Q6GE08 staphylococ
35	793	14.4	335	2	Q5HDC9_STAAC	Q5HDC9 staphylococ
36	793	14.4	335	2	Q7A018_STAAW	Q7A018 staphylococ
37	793	14.4	335	2	Q7A3R9_STAAU	Q7A3R9 staphylococ
38	793	14.4	335	2	Q99RK7_STAAU	Q99RK7 staphylococ
39	778	14.1	321	2	Q4L9U7_STAHJ	Q4L9U7 staphylococ
40	771	14.0	390	2	Q5KY23_GEOKA	Q5KY23 geobacillus
41	767	13.9	391	2	Q749W3_GEOSL	Q749W3 geobacter s
42	763.5	13.8	395	2	Q73KM3_TREDE	Q73KM3 treponema d
43	755	13.7	398	2	Q6KH66_MYCWO	Q6KH66 mycoplasma
44	743.5	13.5	400	2	Q5V3N9_HALMA	Q5V3N9 haloarcula
45	739.5	13.4	392	2	Q67N86_SYMTH	Q67N86 symbiobacte

#### ALIGNMENTS

RESULT 1  
BIOF\_BACSU  
ID BIOF\_BACSU STANDARD; PRT; 389 AA.  
AC P53556;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 13-SEP-2005 (Rel. 48, Last annotation update)  
DE 8-amino-7-oxononanoate synthase (EC 2.3.1.47) (AONS) (8-amino-7-ketopelargonate synthase) (7-keto-8-amino-pelargonic acid synthetase)  
DE (7-KAP synthetase) (L-alanine-piimelyl CoA ligase).  
GN Name=biOf; OrderedLocusNames=BSU30220;  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=96312354; PubMed=8763940;  
RA Bower S., Perkins J.B., Yocum R.R., Howitt C.L., Rahaim P., Pero J.;  
RT "Cloning, sequencing, and characterization of the Bacillus subtilis biotin biosynthetic operon."  
RL J. Bacteriol. 178:4122-4130(1996).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=168;  
RX MEDLINE=98048467; PubMed=9387221;  
RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.-D.;  
RT "Sequencing and functional annotation of the Bacillus subtilis genes in the 200 kb rmb-dnaB region."  
RL Microbiology 143:3431-3441(1997).  
RN [3]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=168;  
RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;  
RA Kunst F., Ogawara N., Mosser I., Albertini A.M., Alloni G., Azevedo V., Betero M.G., Bessieres P., Bolotin A., Borchert S., Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S., Brouillet S., Bruchi C.V., Caldwell B., Capuano V., Carter N.M., Choi S.-K., Codani J.-J., Connerton I.F., Cummings N.J., Daniel R.A., Denizot F., Devine K.M., Duesterhoef A., Ehrlich S.D., Emerson P.T., Entian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,

RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
RA Ghim S.-Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,  
RA Guileppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,  
RA Hilbert H., Holsappel S., Hosono S., Hulio M.-F., Itaya M.,  
RA Jones L.-M., Joris B., Karamata D., Kashahara Y., Klaerr-Bianchard M.,  
RA Klein C., Kobayashi Y., Koetter P., Koningsgatein G., Krogh S.,  
RA Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,  
RA Lazarevic V., Lee S.-M., Levine A., Liu H., Masuda S., Maue C.,  
RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S.,  
RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,  
RA Park S.-H., Farro V., Fohl T.M., Portetelle D., Porwollik S.,  
RA Prescott A.M., Prescan E., Pujic P., Purnelle B., Rapoport G.,  
RA Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,  
RA Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R.,  
RA Scoffone P., Sekiguchi J., Sekowska A., Seror S.J., Serror P.,  
RA Shin B.-S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,  
RA Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P.,  
RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,  
RA Vagstadottir A., Viari A., Wambutt R., Wedler E., Wedler H.,  
RA Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K.,  
RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.-F., Zumstein E.,  
RA Yoshikawa H., Danchin A.,  
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*  
RT *subtilis*,";  
RL Nature 390:249-256(1997).  
CC -I- CATALYTIC ACTIVITY: 6-carboxyhexanoyl-CoA + L-alanine = 8-amino-7-  
CC oxononanoate + CoA + CO(2).  
CC -I- COPACTOR: Pyridoxal phosphate.  
CC -I- PATHWAY: Cofactor biosynthesis; biotin biosynthesis; biotin from  
CC 6-carboxyhexanoyl-CoA: step 1.  
CC -I- SIMILARITY: Belongs to the class-II pyridoxal-phosphate-dependent  
CC aminotransferase family.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
CC EMBL; U51868; AB17459.1; -; Genomic\_DNA.  
CC EMBL; AF008220; AAC00263.1; -; Genomic\_DNA.  
CC EMBL; Z99119; CAB15000.1; -; Genomic\_DNA.  
CC PIR; F69594; F69594.  
CC HSP; P07912; 1FC4.  
CC Subtilist; BG1527; bioF.  
CC InterPro; IPR004839; Aminotrans\_I/II.  
CC InterPro; IPR001917; Aminotrans\_II.  
CC InterPro; IPR004723; BioF.  
CC Pfam; PF00155; Aminotran\_1.2; 1.  
CC TIGRFAMs; TIGR00858; bioF\_1.  
CC PROSITE; PS00599; AA\_TRANSFER\_CLASS\_2; 1.  
KW Biotin biosynthesis; Complete proteome; Pyridoxal phosphate;  
KW Transferrase.  
FT BINDING 237 237 Pyridoxal phosphate (covalent) (By  
FT similarity).  
SQ SEQUENCE 369 AA; 42581 MW; 02B31E16D68F7D6C CRC64;  
Alignment Scores:  
Pred. No.: 4.98e-131 Length: 389  
Score: 1836.00 Matches: 359  
Percent Similarity: 100.0% Conservative: 1  
Best Local Similarity: 99.7% Mismatches: 0  
Query Match: 33.3% Indels: 0  
DB: 1 Gaps: 0  
US-10-681-086-1 (1-3156) x BIOF\_BACSU (1-389)  
QY 22 TTGAAGATTGATTCCTCGTTAAACAGCGGTTAGACAGAAATGAAGAAGCGCGGTACAT 81  
Db :|||||  
Db 1 MetLysIleAspSerTrpLeuAsnGluArgLeuAspArgMetLysGluAlaGlyValHis 20  
QY 82 CGTAACCTCGCGGTCAATGGATGGAGCGCGGTTCCAGAGAGGAATATTGGATGGCGAAAT 141  
|||||

## RESULT 2

Q8KZM9\_BACSU  
ID Q8KZM9\_BACSU PRELIMINARY; PRT; 371 AA.  
AC Q8KZM9;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

Db 21 ArgAsnLeuArgSerMetAspGlyAlaProValProGluArgAsnIleAspGlyGluAsn 40  
QY 142 CAAACGGTCTGGCTCTCAACAATATTATTAGGCGTCGCAAGCGATAGACGTTTCATCGAT 201  
Db |||||||  
Db 41 GluThrValTrpSerSerAsnAsnTrpLeuGlyLeuAlaSerAspArgArgLeuIleAsp 60  
QY 202 GCAGCCCAACACAGCATTTCAGCAATTTGGGACAGAGAACGACGCGTTTCACGTTTAAACGACA 261  
Db |||||||  
Db 61 AlaAlaGluThrAlaLeuGluGlnPheGlyThrGlySerSerGlySerArgLeuThrThr 80  
QY 262 GGCATTTTCGCTCGCATGAAAAAGCTAGAAAGAGATATGCCAGCTTTAACTGACAGAA 321  
Db |||||||  
Db 81 GlyAsnSerValTrpHisGluLysLeuGluLysLysIleAlaSerPheLysLeuThrGlu 100  
QY 322 GCGGCGCTCGTGTTCGAGCGGTACTTCGCCCAATGTCGCGTTCCTTCATCTCTCCCA 381  
Db |||||||  
Db 101 AlaAlaLeuLeuPheSerSerGlyTrpLeuAlaAsnValGlyValLeuSerSerLeuPro 120  
QY 382 GAAAGGAAGATGTCATTTTAAAGTGACAGCTCAATCATCATGCAAGTATGATCGACGGCTGC 441  
Db |||||||  
Db 121 GluLysGluAspValIleLeuSerAspGlnLeuAsnHisAlaSerMetIleAspGlyCys 140  
QY 442 CGACTTTTAAAGCTGATACAGTTGTTTATCGGCATATTGATATGATATGATATGATATG 501  
Db |||||||  
Db 141 ArgLeuSerLysAlaAspThrValValTrpArgHisIleAspMetAsnAspLeuGluAsn 160  
QY 502 AAGCTGATCAACACACAGCGTTTATCAGCGCGTTTATCGTAAACAGCGAGTATTCAGC 561  
Db |||||||  
Db 161 LysLeuAsnGluThrGlnArgTrpGlnArgPheIleValThrAspGlyValPheSer 180  
QY 562 ATGGATGGCAATTCGCCCTCTTGTATCAGATCATCTCACTTGGCAACGCTATCATGCGC 621  
Db |||||||  
Db 181 MetAspGlyThrIleAlaProLeuAspGlnIleSerLeuAlaLysArgTygHisAla 200  
QY 622 TTCGTGGTCTGTGATGATGCCACGCAACAGGAGTTTTCGGCGATTCGGGCAACAGGACG 681  
Db |||||||  
Db 201 PheValValValAspAlaHisAlaThrGlyValLeuGlyAspSerGlyGlnGlyThr 220  
QY 682 AGTGAATACCTTGGTGTTCGCGCGATGTTATCGGCACCTTATCGCACCTTAAACCAAGCTGTGGC 741  
Db |||||||  
Db 221 SerGluTrpPheGlyValCysProAspIleValIleGlyThrLeuSerLysAlaValGly 240  
QY 742 GCGGAGGAGGTTTTCGGCGAGGATCAGCGGCTTTCATCGACTTTTTCGCTGAACCATGCC 801  
Db |||||||  
Db 241 AlaGluGlyGlyPheAlaAlaGlySerAlaValPheIleAspPheLeuLeuAsnHisAla 260  
QY 802 AGAACATTATCTTCAACCGCTATTCGCCGACGACGAGCTGTGCGGCTGCTCAGAGGCT 861  
Db |||||||  
Db 261 ArgThrPheIlePheGluThrAlaIleProProAlaSerCysAlaAlaHisGluAla 280  
QY 862 TTCACATCATCTGAAGCAATATGGGTTATGTGTGTAAGGAGATCATCACACCCATTATT 921  
Db |||||||  
Db 281 PheAsnIleIleGluAlaSerArgGluLysArgGlnLeuLeuPheSerTrpIleSerMet 300  
QY 922 ATCAGAACCATCTGAAGAAATATGGGTTATGTGTGTAAGGAGATCATCACACCCATTATT 981  
Db |||||||  
Db 301 IleArgThrSerLeuLysAsnMetGlyTyTrpValValLysGlyAspHisThrProIleIle 320  
QY 982 CTTGTAGTCATTCGCGATGCCCATATAAACCGTCTCTATTTCCTGCTGAAAACTGACGGCAAG 1041  
Db |||||||  
Db 321 ProValValIleGlyAspAlaHisLysThrValLeuPheAlaGluLysLeuGlnGlyLys 340  
QY 1042 GGAATTTATGCTCTCGCATTCGGCGCCCAACCGTTTCGCGCGGTTGAACCGCGGATTCGA 1101  
Db |||||||  
Db 341 GlyIleTyTrpAlaProAlaIleArgProProThrValAlaProGlyGluSerArgIleArg 360

DE 8-amino-7-oxononanoate synthase.  
GN Name=bioF;  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RA NUCLEOTIDE SEQUENCE.  
RP Sasaki M., Kawamura F., Kurusu Y.;  
RT "Genetic Analysis of an Incomplete bio Operon in a Biotin Auxotrophic Strain of Bacillus subtilis Natto OK2.";  
RL Biosci. Biotechnol. Biochem. 68:739-742(2004).  
DR EMBL; AB080666; BAC03241.1; -; Genomic\_DNA.  
DR HSP; PF07912; IFC4.  
DR GO; GO:0008710; F:8-amino-7-oxononanoate synthase activity; IEA.  
DR GO; GO:0016769; F:transferase activity, transferring nitrogen. . .; IEA.  
DR GO; GO:0009058; P:biosynthesis; IEA.  
DR GO; GO:0009102; P:biotin biosynthesis; IEA.  
DR InterPro; IPR004839; AminoTrans\_I.  
DR InterPro; IPR004917; AminoTrans\_II.  
DR InterPro; IPR004723; BioF.  
DR TIGRFAMs; TIGR00858; bioF\_1.  
DR PROSITE; PS00599; AA\_TRANSFER\_CLASS\_2; UNKNOWN 1.  
SQ SEQUENCE 371 AA; 40446 MW; FDF0204622D226F1 CRC64;  
  
Alignment Scores:  
Pred. No.: 7.9e-119 Length: 371  
Score: 1676.00 Matches: 330  
Percent Similarity: 93.3% Conservative: 6  
Best Local Similarity: 91.7% Mismatches: 6  
Query Match: 30.4% Indels: 18  
DB: 2 Gaps: 1  
  
US-10-681-086-1 (1-3156) x Q8KZM9\_BACSU (1-371)  
QY 22 TTGAAGATGATCTCTGGTTAAACGAGCGGTTAGACAGAAATGAAGAAGCCGCGTACAT 81  
DB 1 MetLysileAspSerTrpLeuAsnAspArgLeuAspIleAlaLysGluAlaGlyValHis 20  
QY 82 CGTAACTCGGTCATGGATGAGCGCGGTCGAGAGGAGATATTGATGCGGAAAT 141  
DB 21 ArgAsnLeuArgSerMetAsnGlyAlaProValProGluArgAsnIleAspGlyGluAsn 40  
QY 142 CAAACGGTCTGGTCTCTCAACAAATATTATGAGGCTCGCAAGCGATAGACGTTTGCAT 201  
DB 41 GlnThrValTrpSerSerAsnAsnTyrLeuGlyLeuAlaSerAspArgArgLeuIleAsp 60  
QY 202 GCAGCCCAACAGCATTTGCAGCAATTTTGGAGCAGGAAGCAGCGGTTCACTTAAACGACA 261  
DB 61 AlaAlaGlnThrAlaLeuGlnGlnPheGlyThrGlySerSerGlySerArgLeuThrThr 80  
QY 262 GGCATTTTCGCTGCGCATGAAGCTAGAAAGAGATTCGCCAGCTTTAACTCAGACGAA 321  
DB 81 GlyAsnSerValTrpHisGlyLysLeuGlyLysIleAlaSerPheLysArgThrGlu 100  
QY 322 GCGCCCTCGTGTTCGAGCGGTTACTTGCACCAATGTCGCGTGTCTTCACTCTGCCA 381  
DB 101 AlaAlaLeuLeuPheSerSerGlyTyrLeuAlaAsnValGlyValLeuSerSerLeuPro 120  
QY 382 GAAAGGAAGATGTCATTTAAAGTGACAGCTCAATCATGCAAGTATGATCGAGCGGTGC 441  
DB 121 GluLysGluAspValIleLeuSerAspGlnLeuAsnHisAlaSerIleIleAspGlyCys 140  
QY 442 CGACTTTCTAAGCTGTATAGTTGTTTATCGCATATTCATGATGATGATCTTGAAC 501  
DB 141 ArgLeuSerLysAlaAspThrValValTyrArgHisIleAspMetAsnAspLeuGluAsn 160  
QY 502 AAGCTGAATGAACACAGCGGTTATCAGCGCGGTTTATCGTAAACAGAGGATTTTCAGC 561  
DB 161 LysLeuAsnGlnThrGlnArgTyrGlnArgArgPheIleValThrAspGlyValPheSer 180  
QY 562 ATGGATGGCAATTCGCCCTCTTTGATCAGATCATCTCATTGCGAAACGCTATCATGCC 621

DB 181 MetAspGlyThrIleValProLeuAspGlnIleIleSerLeuAlaLysArgTyrHisAla 200  
QY 622 TTCTGTCGTCTGTCATGATGCCACGCAACAGAGAGTTTTCGGCGATTCGGCAAGCAACG 681  
DB 201 PheValValValAspAspAlaHisAlaThrGlyValLeuGlyAspSerGlyArgGlyThr 220  
QY 682 AGTGAATACCTTGGTGTTCCTCCGACATGTTTATCGGCACCTTAAGCAAAAGCTGTTGGC 741  
DB 221 SerGluTyrPheGlyValTyrProAspIleValIleGlyThrLeuSerLysAlaValGly 240  
QY 742 GCGGAGGAGGTTTTCGGCAGCATCAGCGGCTCTTCATGACCTTTTGTCTGAACCATGCC 801  
DB 241 ThrGluGlyGlyPheAlaAlaGlySerAlaValPheIleAspPheLeuLeuAsnHisAla 260  
QY 802 AGAACATTTATCTTCAACCGCTATTTCGCCACGACGAGCTGTGCGGTGTCTCAGAGGCT 861  
DB 261 ArgThrPheIlePheGlnThrAlaIleProProIleSerCysAlaAlaHisGluAla 280  
QY 862 TTCAAACATCATTTGAAGCAGCGGAAACACGACAGCTTTTATTTCTTTATATCAGCATG 921  
DB 281 PheAsn----- 282  
QY 922 ATCAGAACCATGCTGAAGATATGCTTATGCTGAAGAGATCACACACCGATTTATT 981  
DB 283 IleArgThrSerLeuLysAsnMetGlyTyrValValLysGlyAspHisThrProIleIle 302  
QY 982 CCTGTAGTCATTCGCGATGCCATATAAACCGTCTCTATTCTGCTGAAAACTCGAGGCAAG 1041  
DB 303 ProValValIleGlyAspAlaHisLysThrValIlePheAlaGluLysLeuGlnGlyLys 322  
QY 1042 GGAATTTATGCTCTCCATTCGCGCCGCAACCGTGTGCGCGGTGAAAGCCGATTCGA 1101  
DB 323 GlyIleTyrAlaProAlaIleArgProProThrValAlaProGlyGluSerArgIleArg 342  
  
RESULT 3  
Q7BU13\_STAAU PRELIMINARY; PRT; 256 AA.  
ID Q7BU13;  
AC Q7BU13;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)  
DE Aminoglycoside adenylyltransferase Aadd (Kanamycin resistance protein)  
DE (O-nucleotidyltransferase (4')).  
GN Name=aadd; Synonyms=aant (4');  
OS Staphylococcus aureus.  
OG Plasmid pSK41, and Plasmid pUB110.  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=1280;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RC PLASMID=pSK41;  
RX MEDLINE=98389645; PubMed=9721269;  
RA Berg T., Firth N., Apisridej S., Hettiaratchi A., Leelaporn A.,  
RA Skurray R.A.;  
RT "Complete nucleotide sequence of pSK41: evolution of staphylococcal  
conjugative multiresistance plasmids.";  
RL J. Bacteriol. 180:4350-4359(1998).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=HUC19; PLASMID=pUB110;  
RX MEDLINE=20316808; PubMed=10858352;  
RX DOI=10.1128/AAC.44.7.1906-1910.2000;  
RA Oliveira D.C., Wu S.W., de Lencastre H.;  
RT "Genetic organization of the downstream region of the mecA element in  
methicillin-resistant Staphylococcus aureus isolates carrying  
different polymorphisms of this region.";  
RL Antimicrob. Agents Chemother. 44:1906-1910(2000).  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=N315;  
RX MEDLINE=92183847; PubMed=1544435; DOI=10.1016/0014-5793(92)80039-J;  
RA Hiramatsu K., Asada K., Suzuki E., Okonogi K., Yokota T.;  
RT "Molecular cloning and nucleotide sequence determination of the

RT regulator region of mecA gene in methicillin-resistant Staphylococcus  
RT aureus (MBSA)";  
RL PERS Lett. 298:133-136 (1992).  
RN [4]  
RC NUCLEOTIDE SEQUENCE.  
RX STRAIN=N315;  
RY MEDLINE=99278010; PubMed=10348769;  
RA Ito T., Katayama Y., Hiramatsu K.;  
RA "Cloning and nucleotide sequence determination of the entire mec DNA  
RT of pre-methicillin-resistant Staphylococcus aureus N315.";  
RL Antimicrob. Agents Chemother. 43:1449-1458 (1999).  
RN [5]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=N315;  
RX MEDLINE=2119321; PubMed=11302791;  
RY DOI=10.1128/AAC.45.5.1323-1336.2001;  
RA Ito T., Katayama Y., Asada K., Mori N., Tsutsumimoto K.;  
RA "Structural comparison of three types of staphylococcal cassette  
RT chromosome mec integrated in the chromosome in methicillin-resistant  
RT Staphylococcus aureus.";  
RL Antimicrob. Agents Chemother. 45:1323-1336 (2001).  
RN [6]  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=N315;  
RX Ito T., Okuma K., Xue M.X., Yuzawa H., Hiramatsu K.;  
RA "Insights on antibiotic resistance of Staphylococcus aureus from its  
RT whole genome: genomic island SCC.";  
RL Drug Resist. Updat. 6:41-52 (2003).  
DR EMBL; AF051917; AAC61950.1; -; Genomic DNA.  
DR EMBL; AF181950; AAF85647.1; -; Genomic DNA.  
DR EMBL; D86934; BAA82229.1; -; Genomic DNA.  
DR SMR; O7BU13; 4-256.  
DR GO; GO:0016779; F:nucleotidyltransferase activity; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR InterPro; IPR002934; DNA\_pol\_beta.  
DR InterPro; IPR012481; KNTase\_C.  
DR Pfam; PF07827; KNTase\_C; 1.  
DR Pfam; PF01909; NTP\_transf\_2; 1.  
DR Pfam; PF01909; NTP\_transf\_2; 1.  
KW Pfam; Transferase.  
SQ SEQUENCE 256 AA; 29166 MW; B45EF8CDBD00654E CRC64;

Alignment Scores:  
Pred. No.: 5,35e-94 Length: 256  
Score: 1350.00 Matches: 256  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 24.5% Indels: 0  
DB: 2 Gaps: 0

US-10-681-086-1 (1-3156) x Q7BU13\_STAAR (1-256)

QY 1570 ATGAGAAATAGTGAATGGACCAATATATGACTAGAGAGAAAGATGATTTTCAT 1629  
Db 1 MetArgIleValAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHis 20  
QY 1630 GAAATTAAGGACGAATATGATGATAATGCGATGATGTTAAGCTATTGTTTAT 1689  
Db 21 GluIleLysGluArgIleLysAspLysIleGlyAspValLysAlaIleGlyValTyr 40  
QY 1690 GGCTCTCTTGGTGCAGACTGATGGCCCTATTTCGGATATTGAGATGATGTTTCATG 1749  
Db 41 GlySerLeuGlyArgGlnThrAspGlyProTyrSerAspIleGluMetCysValMet 60  
QY 1750 TCACAGAGGAGCAGAGATTCAGCCATGATGGACCAACCGGTGAGTGAAGGTGGAAGTG 1809  
Db 61 SerThrGluGluAlaGluPheSerHisGluTrpThrThrGlyGluTrpLysValGluVal 80  
QY 1810 AATTTTGATAGCGAGAGATTTCTACTAGATATGATCTCAGGTGGAATCAGATTGGCGC 1869  
Db 81 AsnPheAspSerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpPro 100  
QY 1870 CTTACACATGGTCAATTTTCTCTATTGTTCCGATTTTATGATTCAGGTGGATCTAGAG 1929  
Db 11 CTTACACATGGTCAATTTTCTCTATTGTTCCGATTTTATGATTCAGGTGGATCTAGAG 1929

Db 101 LeuThrHisGlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGlu 120  
QY 1930 AAGTGTATCAAACTGCTAATCGTAGAAGCCCAACGTTCCACGATGCGATTTGTGCC 1989  
Db 121 LysValTyrGlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAla 140  
QY 1990 CTTATCGTAGAAGAGCTGTTTGAATATGACGAGCAAAATGGCGTAAATATTCGTGTCAAGGA 2049  
Db 141 LeuIleValGluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGly 160  
QY 2050 CCGACAACATTTTACCATCTTCAGTACAGGTAGCAATGGCAGGTGCCATGTTGATT 2109  
Db 161 ProThrThrPheLeuProSerLeuThrValAlaMetAlaGlyAlaMetLeuIle 180  
QY 2110 GGCTGCATCATCGCATCTGTATACGACGAGCGCTTCGGTCTTAACGAAGCAGTTAAG 2169  
Db 181 GlyLeuHisArgIleCysTyrThrSerAlaSerValLeuThrGluAlaValLys 200  
QY 2170 CAATCAGATCTTCCTCAGGTATGACCATCTGTCAGGTTCGTAATGTCGTGTCACACTT 2229  
Db 201 GlnSerAspLeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeu 220  
QY 2230 TCCGACTCTGAGAACTTCCTGGAATCGTAGAATTTTCGGAATGGGATTCAGAGTGG 2289  
Db 221 SerAspSerGluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrp 240  
QY 2290 ACAGAACGACGAGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337  
Db 241 ThrGluArgHisGlyTyrIleValAspValSerLysArgIleProPhe 256

RESULT 4  
Q6GKR2\_STAAR PRELIMINARY; PRT; 256 AA.  
ID Q6GKR2; 2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Kanamycin nucleotidyltransferase (EC 2.7.7.-).  
GN Name=knt; Synonyms=kan; OrderedLocusNames=SAR0033;  
OS Staphylococcus aureus (strain MRSA252).  
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.  
OX NCBI\_TaxID=28458;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX PubMed=15213324; DOI=10.1073/pnas.0402521101;  
RA Holden M.T.G., Feil E.J., Lindsay J.A., Peacock S.J., Day N.P.J.,  
RA Enright M.C., Foster T.J., Moore C.E., Hurst L., Atkin R., Barron A.,  
RA Bason N., Bentley S.D., Chillingworth C., Chillingworth T.,  
RA Churcher C., Clark L., Corton C., Cronin A., Doggett J., Dowd L.,  
RA Feltwell T., Hance Z., Harris B., Hauser H., Holroyd S., Jagels K.,  
RA James K.D., Lennard N., Line A., Mayes R., Moule S., Mungall K.,  
RA Ormond D., Quail M.A., Rabinowitsch E., Rutherford K.M., Sanders M.,  
RA Sharp S., Simmonds M., Stevens K., Whitehead S., Barrell B.G.,  
RA Spratt B.G., Parkhill J.;  
RT "Complete genomes of two clinical Staphylococcus aureus strains:  
RL evidence for the rapid evolution of virulence and drug resistance.";  
RL Proc. Natl. Acad. Sci. U.S.A. 101:9786-9791 (2004).  
DR EMBL; BX571856; CAG39062.1; -; Genomic DNA.  
DR SMR; Q6GKR2; 4-256.  
DR GO; GO:0016779; F:nucleotidyltransferase activity; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR InterPro; IPR002934; DNA\_pol\_beta.  
DR Pfam; PF01909; NTP\_transf\_2; 1.  
KW Complete proteome; Nucleotidyltransferase; Transferase.  
SQ SEQUENCE 256 AA; 29166 MW; B45EF8CDBD00654E CRC64;

Alignment Scores:  
Pred. No.: 5,35e-94 Length: 256  
Score: 1350.00 Matches: 256  
Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 24.5% Indels: 0  
DB: 2 Gaps: 0

US-10-681-086-1 (1-3156) x Q6GKR2\_STAAR (1-256)

QY 1570 ATGAGATAGTGAATGACCAATAATACTAGTAGAGAAAGAAATGAAGATTGTTTCAT 1629

Db 1 MetArgileValAsnGlyProIleMetThrArgGluGluArgMetLysIleValHis 20

QY 1630 GAAATTAAGGAACAAATATTGGATAATATGGGATGATGTTAAGGCTATTGGTGTATTAT 1689

Db 21 GluilelsgluargileLeuAspLysTyThrGlyAspAspValLysAlaIleGlyValTy 40

QY 1690 GGCTCTCTGGTCTCAGACTGATGGCCCTATTTCGGATATTTCGAGATATTGAGATGTTGTCATG 1749

Db 41 GlySerLeuGlyArgGlnThrAspGlyProTyThrSerAspIleGluMetMetCysValMet 60

QY 1750 TCACACAGACAGACAGTTTCACCCATGAATGACACACCGGTGAGTGGAGGTGGAGTG 1809

Db 61 SerThrGluGluAlaGluPheSerHisGluTrpThrThrGlyGluTrpLysValGluVal 80

QY 1810 AATTTTATGATGACGAGAGATTTCTACTAGATTATGATCTCAGGTGGAATCAGATTGGCCG 1869

Db 81 AsnPheAspSerGluGluIleLeuLeuAspTyThrAlaSerGlnValGluSerAspTrpPro 100

QY 1870 CTTCACATCGTCAATTTTCTTATTTTCCGATTTTATGATTCAGTGGATCTTAGAG 1929

Db 101 LeuThrHisGlyGlnPheSerIleLeuProIleTyThrAspSerGlyGlyTyThrLeuGlu 120

QY 1930 AAGTGTATCAACTGTAATCGGTAGAACCCCAACCGTTCACGATGCGATTGGCC 1989

Db 121 LysValTyThrGlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAla 140

QY 1990 CTTATCTGACAGACGCTTTTGAATATGACGCAATGGCGTAATATTTCGTGTGCAAGGA 2049

Db 141 LeuileValGluGluLeuPheGluTyThrAlaGlyLysTrpArgAsnIleArgValGlnGly 160

QY 2050 CCGACCAACTTCTACCATCTCTGAGTACAGTAGCAATTCGGAATGGGATTTCAGGAGTGG 2109

Db 161 ProThrThrPheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuile 180

QY 2110 GGCTGTCATCATGTCATCTGTTATACGACGAGCGCTTCGGTCTTAACCTGAAGCAGTTAAG 2169

Db 181 GlyLeuHisArgileCysTyThrThrSerAlaSerValLeuThrGluAlaValLys 200

QY 2170 CAATCAGATCTTCTTCAGGCTTATGACCTCTGCGCAGTTCGTAATGCTGGTCAACTT 2229

Db 201 GlnSerAspLeuProSerGlyTyThrAspHisLeuValMetSerGlyGlnLeu 220

QY 2230 TCCGACTCTGAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTTCAGGAGTGG 2289

Db 221 SerAspSerGluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrp 240

QY 2290 ACAGACGACGAGATATAGTGGATGTGTCAAAACGATACCATTT 2337

Db 241 ThrGluArgHisGlyTyThrIleValAspValSerLysArgileProPhe 256

## RESULT 5

O87369\_STAAR PRELIMINARY; PRT; 256 AA.

AC O87369;

DT 01-NOV-1998 (TrEMBLrel. 08, Created)

DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Kanamycin nucleotidyltransferase.

GN Name=aadB; OrderedLocusNames=SAV0035;

OS Staphylococcus aureus (strain Mu50 / ATCC 700699).

OC Bacteria; Firmicutes; Bacillales; Staphylococcus.

OX NCBI\_TaxID=158878;

RN [1]

RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].

RX MEDLINE=21311952; PubMed=11418146; DOI=10.1016/S0140-6736(00)04403-2;

RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,

RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,

RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Hosoyama A.,

RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,

RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,

RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,

RA Hattori M., Ogasawara N., Hayaishi H., Hiramatsu K.,

RT "Whole genome sequencing of methicillin-resistant Staphylococcus

RL Lancet 357:1225-1240 (2001).

DR EMBL; BA000017; BAB56197.1; -; Genomic\_DNA.

DR PIR; S09565; S09565.

DR HSP; P05057; IKNY.

DR SNR; O87369; 4-256.

DR GO; GO:0016779; F:nucleotidyltransferase activity; IEA.

DR GO; GO:0016740; F:transferase activity; IEA.

DR InterPro; IPR02934; DNA\_pol\_beta.

DR Pfam; PF01909; NTP\_transf\_2; 1.

KW Complete proteome; Transferase.

SQ SEQUENCE 256 AA; 29166 MW; B45EF8CDBD00654E CRC64;

## Alignment Scores:

Pred. No.:	5,35e-94	Length:	256
Score:	1350.00	Matches:	256
Percent Similarity:	100.0%	Conservative:	0
Best Local Similarity:	100.0%	Mismatches:	0
Query Match:	24.5%	Indels:	0
DB:	2	Gaps:	0

US-10-681-086-1 (1-3156) x O87369\_STAAR (1-256)

QY 1570 ATGAGATAGTGAATGACCAATAATACTAGTAGAGAAAGAAATGAAGATTGTTTCAT 1629

Db 1 MetArgileValAsnGlyProIleMetThrArgGluGluArgMetLysIleValHis 20

QY 1630 GAAATTAAGGAACAAATATTGGATAATATGGGATGATGTTAAGGCTATTGGTGTATTAT 1689

Db 21 GluilelsgluargileLeuAspLysTyThrGlyAspAspValLysAlaIleGlyValTy 40

QY 1690 GGCTCTCTGGTCTCAGACTGATGGCCCTATTTCGGATATTTCGAGATGTTGTCATG 1749

Db 41 GlySerLeuGlyArgGlnThrAspGlyProTyThrSerAspIleGluMetCysValMet 60

QY 1750 TCACACAGACAGACAGTTTCACCCATGAATGACACACCGGTGAGTGGAGGTGGAGTG 1809

Db 61 SerThrGluGluAlaGluPheSerHisGluTrpThrThrGlyGluTrpLysValGluVal 80

QY 1810 AATTTTATGATGACGAGATTTCTACTAGATTATGATCTCAGGTGGAATCAGATTGGCCG 1869

Db 81 AsnPheAspSerGluGluIleLeuLeuAspTyThrAlaSerGlnValGluSerAspTrpPro 100

QY 1870 CTTCACATCGTCAATTTTCTTATTTTCCGATTTTATGATTCAGTGGATCTTAGAG 1929

Db 101 LeuThrHisGlyGlnPheSerIleLeuProIleTyThrAspSerGlyGlyTyThrLeuGlu 120

QY 1930 AAGTGTATCAACTGTAATCGGTAGAACCCCAACCGTTCACGATGCGATTGGTGGCC 1989

Db 121 LysValTyThrGlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAla 140

QY 1990 CTTATCTGACAGACGCTTTTGAATATGACGCAATGGCGTAATATTTCGTGTGCAAGGA 2049

Db 141 LeuileValGluGluLeuPheGluTyThrAlaGlyLysTrpArgAsnIleArgValGlnGly 160

QY 2050 CCGACCAACTTCTACCATCTCTGAGTACAGTAGCAATTCGGAATGGGATTTCAGGAGTGG 2109

Db 161 ProThrThrPheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuile 180

QY 2110 GGCTGTCATCATGTCATCTGTTATACGACGAGCGCTTCGGTCTTAACCTGAAGCAGTTAAG 2169

Db 181 GlyLeuHisArgileCysTyThrThrSerAlaSerValLeuThrGluAlaValLys 200

QY 2170 CAATCAGATCTTCTTCAGGCTTATGACCTCTGCGCAGTTCGTAATGCTGGTCAACTT 2229

Db 201 GlnSerAspLeuProSerGlyTyThrAspHisLeuValMetSerGlyGlnLeu 220

QY 2230 TCCGACTCTGAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTTCAGGAGTGG 2289

```
|||||
Db 221 SerAspSerGluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrp 240
QY 2290 ACAGAACGACGCGATATATAGTGGTGTGCAAAACGCATACCATTT 2337
Db 241 ThrGluArgHisGlyIleValAspValSerLysArgIleProPhe 256
RESULT 6
Q7A8D0_STAA0 PRELIMINARY; PRT; 256 AA.
AC Q7A8D0;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Kanamycin nucleotidyltransferase
GN Name=adD; OrderedLocuNames=SA0033;
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RA MEDLINE=21311952; PubMed=11418146; DOI=10.1016/S0140-6736(00)04403-2;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Maruyama A., Murakami H., Haseyama A.,
RA Mizutani-Ui Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of meticillin-resistant Staphylococcus
RT aureus.";
RL Lancet 357:1225-1240(2001).
DR EMBL; BA000018; BAB41251.1; -; Genomic_DNA.
DR SMR; Q7A8D0; 4-256.
DR GO; GO:0016779; F:nucleotidyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR002934; DNA pol beta.
DR Pfam; PF01909; NTP transf 2_1.
KW Complete proteome; Transferase.
SQ SEQUENCE 256 AA; 29166 MW; B45E98CDBD00654E CRC64;

Alignment Scores:
Pred. No.: 5,35e-94 Length: 256
Score: 1350.00 Matches: 256
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 24.5% Indels: 0
DB: 2 Gaps: 0

US-10-681-086-1 (1-3156) x Q7A8D0_STAA0 (1-256)
QY 1570 ATGAGATAGTGAATGGACCAATATATAGTAGAGAGAGAAAGATGATGTTTCAT 1629
Db 1 MetArgIleValAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHis 20
QY 1630 GAATTAAGGACGAATATCGATAAATCGGATGATGTTAAGGCTATTGTTTAT 1689
Db 21 GluIleLysGluArgIleLeuAspLysTyrGlyAspValLysAlaIleLysValTyr 40
QY 1690 GGCTCTCTGTGTCGACACTGATGGGCCCTATTTCGGATATTGAGATGATGTTGTCATG 1749
Db 41 GlySerLeuGlyArgGlnThrAspGlyProTyrSerAspIleGluMetCysValMet 60
QY 1750 TCAACAGAGGAGCAGAGTTGACCCATGAATGGACACCGGTGAGTGAAGGTGGAAGTG 1809
Db 61 SerThrGluGluAlaGluPheSerHisGluTrpThrThrGlyGluTrpLysValGluVal 80
QY 1810 AATTTTGATACGAGAGATTTCTACTAGATATGCTCATCGAGTGGATTCAGATTGGCGG 1869
Db 81 AsnPheAspSerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpPro 100
QY 1870 CTTACACATGGTCAATTTTCTCTATTTCGCGATTATGATTCAGGTGGATCTTAGAG 1929
|||||

101 LeuThrHisGlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGlu 120
QY 1930 AAAGTGTATCAAACTGCTAAATCGTAGAAGCCCAACGTTTCACGATGCGATTGTGQC 1989
Db 121 LysValTyrGlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAla 140
QY 1990 CTTATCGTAGAAGAGCTGTTTGAATATGACAGGCAATGGCGTAATATTCGTGTGCAAGGA 2049
Db 141 LeuIleValGluGluLeuPheGluTyrAlaGlyLysTyrArgAsnIleArgValGlnGly 160
QY 2050 CCGACAAACATTTCTACCATCTTGCAGTGTACAGTAGCAATGGCAGTGGCATGTTGATT 2109
Db 161 ProThrThrPheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIle 180
QY 2110 GGTCGTGCATCATCGCATCTGTATACAGCAGCGCTTCGGTCTTAACTGAAGCAGTTAAG 2169
Db 181 GlyLeuHisHisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLys 200
QY 2170 CAATCAGATCTTCTTCAGGTTATGACCATCTGTGCGCAGTTCGTAATGTCGTGCACTT 2229
Db 201 GlnSerAspLeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeu 220
QY 2230 TCCGACTCTGAGAACTTCTCGAATCGCTAGAGAAATTTCTGGAATGCGATTCAGAGTGG 2289
Db 221 SerAspSerGluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrp 240
QY 2290 ACAGAACGACGCGATATATAGTGGATGTCGTAACAAACGCATACCATTT 2337
Db 241 ThrGluArgHisGlyIleValAspValSerLysArgIleProPhe 256

RESULT 7
KANU_STAA0
ID_KANU_STAA0 STANDARD; PRT; 253 AA.
AC P05057;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Kanamycin nucleotidyltransferase (EC 2.7.7.-) (Neo(R)).
GN Name=knt; Synonyms=kan;
OS Staphylococcus aureus.
OG Plasmid pUB110.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP NUCLEOTIDE SEQUENCE
RA MEDLINE=86174350; PubMed=3007933; DOI=10.1007/BF00330534;
RA Mueller R.E., Ano T., Imanaka T., Aiba S.;
RT "Complete nucleotide sequences of Bacillus plasmids pUB110dB, pRBH1
RT and its copy mutants.";
RL Mol. Gen. Genet. 202:169-171(1986).
RN [2]
RP NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 1-7.
RX MEDLINE=8506820; PubMed=6090428;
RA Matsumura M., Katakura Y., Imanaka T., Aiba S.;
RT "Enzymatic and nucleotide sequence studies of a kanamycin-inactivating
RT enzyme encoded by a plasmid from thermophilic bacilli in comparison
RT with that encoded by plasmid pUB110.";
RL J. Bacteriol. 160:413-420(1984).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=86301853; PubMed=3744038;
RA Bashkurov V.I., Mil'china N.V., Prozorov A.A.;
RT "Nucleotide sequence and physical map of kanamycin-resistant plasmid
RT pUB110 from Staphylococcus aureus.";
RL Genetika 22:1081-1092(1986).
RN [4]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=86206447; PubMed=3010356;
RA McKenzie T., Hoshino T., Tanaka T., Sueoka N.;
RT "The nucleotide sequence of pUB110: some salient features in relation
RT to replication and its regulation.";
RL Plasmid 15:93-103(1986).
RN [5]
```







DR Pfam; PF01909; NTP\_transf\_2; 1.  
KW Antibiotic resistance; Direct protein sequencing; Plasmid;  
KW Transferase;  
SQ SEQUENCE 253 AA; 28825 MW; 9CA14603B2B5DC6 CRC64;

Alignment Scores:  
Pred. No.: 3,03e-92 Length: 253  
Score: 1327.00 Matches: 251  
Percent Similarity: 99.6% Conservative: 1  
Best Local Similarity: 99.2% Mismatches: 1  
Query Match: 24.1% Indels: 0  
DB: 1 Gaps: 0

US-10-681-086-1 (1-3156) x KANU\_BACSP (1-253)

Qy 1579 GTGAATGACCAATAATAATGACTAGAGAGAAAGAAATGAAGATTGTTTCATGAATTAAG 1638  
:::|||||  
Db 1 MethanoglyProfilleMetThrArgGluGluArgMetIysIleValHisGluIleLys 20

Qy 1639 GAACGAATATGGATAAATATGGGATGATGTTAAAGCTATTGCTGTTATGGCTCTCTT 1698  
|||||  
Db 21 GluArgileLeuAspLysTyrGlyAspValIysAlaIleGlyValTyrGlySerLeu 40

Qy 1699 GGTGCTCAGACTGATGGCCCTATTCCGATATTGAGATGATGTTGTCATGTCACAGAG 1758  
|||||  
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetCysValMetSerThrGlu 60

Qy 1759 GAACGACAGATTTCAGCCATGAATGACACACCGTCCAGTGGAGGTGAAGTGAATTTTAT 1818  
|||||  
Db 61 GluAlaGluPheSerHisGluTrpThrGlyGluTrpLysValGluValAsnPheAsp 80

Qy 1819 AGCGAAGAGATTCTACTAGATTATGATTCAGTGGAAATCAGATTGCGCGCTTACACAT 1878  
|||||  
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpProLeuThrHis 100

Qy 1879 GGTCAATTTTCTCTATTGTCGCAATTATGATTTCAGTGGATCTTAGAGAAAGTGTAT 1938  
|||||  
Db 101 GlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120

Qy 1939 CAAACTGCTAAATCGGTAGAACCCAAACGTTCCACGATGCGATTTCGTGCCCTTATCGTA 1998  
|||||  
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140

Qy 1999 GAAGAGCTGTTGTAATATGACGACCAATGCGTAATATTCGTGTGCAAGGACCGACAACA 2058  
|||||  
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160

Qy 2059 TTCTACATCTTGTACTGTACAGGTAGCAATGCGAGTGCATGTCATGTTGATGGTCTGCAT 2118  
|||||  
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180

Qy 2119 CATCGCATCTGTTATACGACGAGCGCTTCGTCTTACTGAAGCAGTTAAGCAATCAGAT 2178  
|||||  
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200

Qy 2179 CTTCCTTCAGGTTATGACCATCTGTGCGAGTTCGTAATGCTGTGTCACACTTCGCACTCT 2238  
|||||  
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220

Qy 2239 GAGAACTTCGGAATCGCTAGAGAAATTCGGAATGGGATTCAGGAGTGGACAGACGA 2298  
|||||  
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240

Qy 2299 CACGGATATATAGTGGATGTCACAAACGATACCATTT 2337  
|||||  
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 10  
Q57514\_BACST  
ID Q57514\_BACST PRELIMINARY; PRT; 253 AA.  
AC Q57514\_1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)

DT 01-FEB-2005 (TrEMBLrel. 29, Last annotation update)  
DE Kanamycin nucleotidyltransferase.  
OS Bacillus stearothermophilus.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Geobacillus.  
OX NCBI\_TaxID=1422;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RA Narumi I., Nakayama N., Nakamoto S., Kihara H.;  
RT "Complete nucleotide sequence of pSTK1, a cryptic plasmid from  
Bacillus stearothermophilus TK015";  
RL Biotechnol. Lett. 15:1013-1016(1993).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RA Narumi I., Nakayama N., Nakamoto S., Kihara H.;  
RT "Bacillus stearothermophilus plasmid pSTK1 replicon is functional in  
Escherichia coli.";  
RL Biotechnol. Lett. 17:475-480(1995).  
RN [3]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=86120983; PubMed=3003740;  
RA Liao H., McKenzie T., Hageman R.;  
RT "Isolation of a thermostable enzyme variant by cloning and selection  
in a thermophile.";  
RL Proc. Natl. Acad. Sci. U.S.A. 83:576-580(1986).  
RN [4]  
RP NUCLEOTIDE SEQUENCE.  
RA Narumi I., Nakayama N., Nakamoto S., Kimura T., Yanagkisawa T.,  
RT Kihara H.;  
RT "Construction of a new shuttle vector pSTE33 and its stabilities in  
Bacillus stearothermophilus, Bacillus subtilis, and Escherichia  
coli.";  
RL Biotechnol. Lett. 15:815-820(1993).  
DR EMBL; D29979; BAA06249.1; -; Genomic DNA.  
DR EMBL; D29978; BAA06245.1; -; Genomic DNA.  
DR HSSP; P05057; 1KNY.  
DR SMR; Q57514; 1-253.  
DR GO; GO:0016779; F:nucleotidyltransferase activity; IEA.  
DR GO; GO:0016740; F:transferase activity; IEA.  
DR InterPro; IPR002934; DNA\_pol\_beta.  
DR Pfam; PF01909; NTP\_transf\_2; 1.  
KW Transferase.  
SQ SEQUENCE 253 AA; 28873 MW; FE573F27646F281E CRC64;

Alignment Scores:  
Pred. No.: 1,47e-91 Length: 253  
Score: 1318.00 Matches: 250  
Percent Similarity: 99.2% Conservative: 1  
Best Local Similarity: 98.8% Mismatches: 2  
Query Match: 23.9% Indels: 0  
DB: 2 Gaps: 0

US-10-681-086-1 (1-3156) x Q57514\_BACST (1-253)

Qy 1579 GTGAATGACCAATAATAATGACTAGAGAGAAAGAAATGAAGATTGTTTCATGAATTAAG 1638  
:::|||||  
Db 1 MetAsnGlyProfilleMetThrArgGluGluArgMetIysIleValHisGluIleLys 20

Qy 1639 GAACGAATATGGATAAATATGGGATGATGTTAAAGCTATTGCTGTTATGGCTCTCTT 1698  
|||||  
Db 21 GluArgileLeuAspLysTyrGlyAspValIysAlaIleGlyValTyrGlySerLeu 40

Qy 1699 GGTGCTCAGACTGATGGCCCTATTCCGATATTGAGATGATGTTGTCATGTCACAGAG 1758  
|||||  
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetCysValMetSerThrGlu 60

Qy 1759 GAACGACAGATTTCAGCCATGAATGACACACCGTCCAGTGGAGGTGAAGTGAATTTTAT 1818  
|||||  
Db 61 GluAlaGluPheSerHisGluTrpThrGlyGluTrpLysValGluValAsnPheTyr 80

Qy 1819 AGCGAAGAGATTCTACTAGATTATGATTCAGTGGAAATCAGATTGCGCGCTTACACAT 1878  
|||||  
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpProLeuThrHis 100

Qy 1879 GGTCAATTTTCTCTATTTCCTGATTTATGATTCAGGTGGATATCTAGAGAAAGTGTAT 1938  
Db |||||GlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120  
Qy 1939 CAAACTGCTAAATCGGTAGAGCCCAACGTTCCACGATGCGATTTGTGCCCTTATCGTA 1998  
Db |||||GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140  
Qy 1999 GAAGAGCTGTTGTAATGATGCGAGCAATATGGCGTATATCTGTGTCGAAGGCCCAACA 2058  
Db |||||GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160  
Qy 2059 TTTCTACCATCTTTCAGCTGTCAGGTAGCAGTGCAGTGCATGTTGATTTGCTGTCAT 2118  
Db |||||PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180  
Qy 2119 CATCGCATCTCTTATACGACGAGCGCTTCGCTTAACTGAAGCAGTTAAAGCAATCAGAT 2178  
Db |||||HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200  
Qy 2179 CTTCTCTCAGGTATGACCATCTGCGCAGTTCGTATATGTCGTCAACTTTCGACTCT 2238  
Db |||||LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220  
Qy 2239 GAGAACTTCTGGAATCGTACAGATTTCTGGAATGGATTTCAGGAGTGCAGACAGACGA 2298  
Db |||||GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGlyTrpThrGluArg 240  
Qy 2299 CACGGATATATAGTGGATGTGTCAAAACGCGATACCATTT 2337  
Db |||||HisGlyTyrIleValAspValSerLysArgIleProPhe 253

## RESULT 11

BIOS BACSU STANDARD; PRT; 335 AA.  
AC P53557;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 13-SEP-2005 (Rel. 46, Last annotation update)  
DE Biotin synthase (EC 2.8.1.6) (Biotin synthetase).  
GN Name=biOB; OrderedLocusNames=BSU30200;  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP NUCLEOTIDE SEQUENCE.  
RX MEDLINE=96312354; PubMed=8763940;  
RA Bower S., Perkins J.B., Yocum R.R., Howitt C.L., Rahaim P., Pero J.;  
RT "Cloning, sequencing, and characterization of the Bacillus subtilis  
RT biotin biosynthetic operon."  
RL J. Bacteriol. 178:4122-4130(1996).  
RN [2]  
RP NUCLEOTIDE SEQUENCE.  
RX STRAIN=168;  
RC MEDLINE=98048467; PubMed=9387221;  
RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.-D.;  
RT "Sequencing and functional annotation of the Bacillus subtilis genes  
RT in the 200 kb rnb-dnaB region."  
RL Microbiology 143:3431-3441(1997).  
RN [3]  
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].  
RC STRAIN=168;  
RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;  
RA Kunst F., Ogasawara N., Moser I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
RA Borries R., Bourcier L., Brans A., Braun M., Brignell S.C., Bron S.,  
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
RA Choi S.-K., Codani J.-C., Connerton I.F., Cummings N.J., Daniel R.A.,  
RA Denizot F., Devine K.M., Duesterhoeft A., Ehrlich S.D., Emmerson P.T.,  
RA Entian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,  
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
RA Ghim S.-Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
RA Guiseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,

RA Hilbert H., Holsappel S., Hosono S., Hullo M.-F., Itaya M.,  
RA Jones L.-W., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M.,  
RA Klein C., Kobayashi Y., Koetter P., Konigstein G., Krogh S.,  
RA Kumano M., Kurica K., Lapidus A., Lardinois S., Lauber J.,  
RA Lazarevic V., Lee S.-M., Levine A., Liu H., Masuda S., Mauel C.,  
RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moesti D., Nakai S.,  
RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,  
RA Park S.-H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,  
RA Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G.,  
RA Rey M., Reynolds S., Rieger M., Rivoita C., Rocha E., Roche B.,  
RA Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R.,  
RA Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Seror P.,  
RA Shin B.-S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,  
RA Takemaru K., Takeuchi M., Takakoshi A., Tanaka T., Terptrak P.,  
RA Tognoni A., Toato V., Uchiyama S., Vandenbol M., Vannier F.,  
RA Vassaroni A., Viari A., Wambutt R., Wedler E., Wedler H.,  
RA Weitzneger T., Winters P., Wipat A., Yamamoto H., Yamane K.,  
RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.-F., Zumstein E.,  
RA Yoshikawa H., Danchin A.;  
RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
RT subtilis";  
RL Nature 390:249-256(1997).  
CC -1- CATALYTIC ACTIVITY: Dethiobiotin + sulfur = biotin.  
CC -1- COFACTOR: Binds 1 4Fe-4S cluster coordinated with 3 cysteines and  
CC an exchangeable S-adenosyl-L-methionine (By similarity).  
CC -1- COFACTOR: Binds 1 2Fe-2S cluster coordinated with 3 cysteines and  
CC 1 arginine (By similarity).  
CC -1- PATHWAY: Cofactor biosynthesis; biotin biosynthesis; biotin from  
CC 6-carboxyhexanoyl-CoA: step 4 [final step].  
CC -1- SIMILARITY: Belongs to the biotin and lipoic acid synthetases  
CC family.  
CC -----  
CC This Swiss-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use as long as its content is in no way modified and this statement is not  
CC removed.  
CC -----  
CC EMBL: U51868; AAB17461.1; -; Genomic DNA.  
CC EMBL: AF008220; AAC00265.1; -; Genomic DNA.  
CC EMBL: Z99119; CAB14998.1; -; Genomic DNA.  
CC PIR: D69594; D69594.  
CC Subtilist; BG11525; bioB.  
CC InterPro: IPR010722; BATS.  
CC InterPro: IPR002684; Biotin synth.  
CC InterPro: IPR006638; Elp3/Miab/Nifb.  
CC InterPro: IPR007197; Radical SAM.  
CC Pfam: PF06968; BATS; 1.  
CC PIRSF: PIRSF001619; Radical SAM; 1.  
CC SMART: SM00729; Elp3; 1.  
CC TIGRFAMS: TIGR00433; bioB; 1.  
KW 2Fe-2S; 4Fe-4S; Biotin biosynthesis; Complete proteome; Iron;  
KW Iron-sulfur; Metal-binding; Transference.  
FT METAL 65 65 Iron-sulfur 1 (4Fe-4S-S-AdoMet) (By  
FT similarity).  
FT METAL 69 69 Iron-sulfur 1 (4Fe-4S-S-AdoMet) (By  
FT similarity).  
FT METAL 72 72 Iron-sulfur 1 (4Fe-4S-S-AdoMet) (By  
FT similarity).  
FT METAL 109 109 Iron-sulfur 2 (2Fe-2S) (By similarity).  
FT METAL 141 141 Iron-sulfur 2 (2Fe-2S) (By similarity).  
FT METAL 201 201 Iron-sulfur 2 (2Fe-2S) (By similarity).  
FT METAL 271 271 Iron-sulfur 2 (2Fe-2S) (By similarity).  
SQ SEQUENCE 335 AA; 36927 MW; D4B2E1A53271ED26 CRC64;  
Alignment Scores:  
Pred. No.: 1,87E-91 Length: 335  
Score: 1337.00 Matches: 263  
Percent Similarity: 92.0% Conservatives: 2  
Best Local Similarity: 91.3% Mismatches: 5  
Query Match: 23.9% Indels: 18  
DB: 1 Gaps: 1

US-10-681-086-1 (1-3156) x BIOB\_BACSU (1-335)

QY 2307 TATAGTGGATGTGCAAAACCCATACCATTTTGAAT----- 2342  
Db 48 TyrGlyLysLysValLysLeuAsnMetIleMetAsnAlaLysSerGlyLeuCysProGlu 67  
QY 2343 -----TCGAAAGCGCGATTTGATCTTACCGGATG 2372  
Db 68 AsnCysGlyTyrCysSerGlnSerAlaIleSerLysAlaProIleGluSerTyrArgMet 87  
QY 2373 GTGAATAAGGAAACGCTGTTGAAGCGCGGCGGCGACGATCTGAATATCGGCACA 2432  
Db 88 ValAsnLysGluThrLeuLeuGluGlyAlaLysArgAlaHisAspLeuAsnIleGlyThr 107  
QY 2433 TATTGTATCGTGGCAAGCGGACAGAGTCCGCTTAACAGAGAAGTGGATCGGTAGAT 2492  
Db 108 TyrCysIleValAlaSerGlyArgGlyProSerAsnArgGluValAspGlnValValAsp 127  
QY 2493 GCGGTTACGAGAAATTAAGAGACGATGCGTCTTGAAGTTCGATGCTTTGGACTGTG 2552  
Db 128 AlaValGlnGluIleLysGluThrTyrGlyLeuLysIleCysAlaCysLeuGlyLeuLeu 147  
QY 2553 AAGCCAGACGAGCGGCGGCTCAAGAGTACGAGGAGTACCGCTATATCAATAATTG 2612  
Db 148 LysProGlnGlnAlaLysArgLeuLysAspAlaGlyValAspArgTyrAsnHisAsnLeu 167  
QY 2613 AATACGTCACAGAGAAACCATTTCAACATCACACCTCACATACATACAGTACAGAGTC 2672  
Db 168 AsnThrSerGlnArgAsnHisSerAsnIleThrThrSerHisThrTyrAspAspArgVal 187  
QY 2673 AATACGTTGAATCGCAAAAGAAATCGGGGCTGTCTCGGCTTACCGGCGCCATTCGGG 2732  
Db 188 AsnThrValGluIleAlaLysGluSerGlyLeuSerProCysSerGlyAlaIleIleGly 207  
QY 2733 ATGAAGGAGACGAAACAGGATGTCAATGACATCCGCAAAAGCTTGAAGCTCTTGAAGCG 2792  
Db 208 MetLysGluThrLysGlnAspValIleAspIleAlaLysSerLeuLysAlaLeuAspAla 227  
QY 2793 GATTCCTTCCTGGAATTTTTCATCGCATTCATCGCAGCGCTTGAAGCGGTAGAGCGTCAAC 2852  
Db 228 AspSerIleProValAsnPheLeuHisAlaIleAspGlyThrProLeuGluGlyValAsn 247  
QY 2853 GAATTAACCCGCTGATTTGTTTAAAGTGTGCGGCTGTTCCGCTTTTATCAATCCATCA 2912  
Db 248 GluLeuAsnProLeuTyrCysLeuLysValLeuAlaLeuPheArgPheIleAsnProSer 267  
QY 2913 AAAGAAATTCGATTTCCGAGGAGAGAGTCAATCTCCGCAATTCGACGATTTAGG 2972  
Db 268 LysGluIleArgIleSerGlyArgGluValAsnLeuArgThrLeuGlnProLeuGly 287  
QY 2973 CTTTACGCGCAAACTCCATTTTGTGCGAGACTACTTAACTCCGCGGCGGAGAGGAG 3032  
Db 288 LeuTyrAlaAlaAsnSerIlePheValGlyAspTyrLeuThrThrAlaGlyGlnGluGlu 307  
QY 3033 ACGGAGATCATAAATGCTGAGTGATTTAGGCTTTGAAGTTCGATCAGTCCGAGAAATG 3092  
Db 308 ThrGluAspHisLysMetLeuSerAspLeuGlyPheGluValGluSerValGluGluMet 327  
QY 3093 AAGCTGATTTAAGTCGAAAGC 3116  
Db 328 LysAlaSerLeuSerAlaLysSer 335

RESULT 12  
Q8KZM7\_BACSU PRELIMINARY; PRT; 335 AA.  
AC Q8KZM7;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Biotin synthetase.  
GN Name=biob;  
OS Bacillus subtilis.

Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
NCBI\_TaxID=1423;  
[1]  
NUCLEOTIDE SEQUENCE.  
Saaki M., Kawamura F., Kuruu Y.;  
"Genetic Analysis of an incomplete bio Operon in a Biotin Auxotrophic  
Strain of Bacillus subtilis Natto OK2.";  
Biosci. Biotechnol. Biochem. 68:739-742(2004).  
-I- CATALYTIC ACTIVITY: Dethiobiotin + sulfur = biotin.  
-I- COFACTOR: Binds a 4Fe-4S cluster coordinated with 3 cysteines and  
an exchangeable S-adenosyl-L-methionine, and a 2Fe-2S cluster  
coordinated with 3 cysteines and 1 arginine (By similarity).  
-I- PATHWAY: Biotin biosynthesis; last step.  
EMBL; AB088066; BAC03243.1; -; Genomic\_DNA.  
HSP; P12996; IR30.  
DR GO:0004076; P:biotin synthase activity; IEA.  
DR GO:0005506; P:iron ion binding; IEA.  
DR GO:0016740; P:transferase activity; IEA.  
DR GO:0009102; P:biotin biosynthesis; IEA.  
DR InterPro: IPR010722; BATS.  
DR InterPro: IPR002684; Biotin synth.  
DR InterPro: IPR007197; Radical SAM.  
DR Pfam: PF06968; BATS; 1.  
DR Pfam: PF04055; Radical SAM; 1.  
DR PIRSF: PIRSF001619; Biotin synth; 1.  
DR SMART: SM00729; Elp3; 1.  
DR TIGRFAMs: TIGR00433; bioB; 1.  
DR 2Fe-2S; 4Fe-4S; Biotin biosynthesis; Iron; Iron-sulfur; Metal-binding;  
KW Transferase.  
SQ SEQUENCE 335 AA; 36880 MW; 82683C6E168D9FAC CRC64;

Alignment Scores:  
Pred. No.: 9,08e-91 Length: 335  
Score: 1308.00 Matches: 260  
Percent Similarity: 92.0% Conservative: 5  
Best Local Similarity: 90.3% Mismatches: 5  
Query Match: 23.7% Indels: 18  
DB: 2 Gaps: 1

US-10-681-086-1 (1-3156) x Q8KZM7\_BACSU (1-335)

QY 2307 TATAGTGGATGTGCAAAACCCATACCATTTTGAAT----- 2342  
Db 48 TyrGlyLysLysValLysLeuAsnMetIleMetAsnAlaLysSerGlyLeuCysProGlu 67  
QY 2343 -----TCGAAAGCGCGATTTGATCTTACCGGATG 2372  
Db 68 AsnCysGlyTyrCysSerGlnSerAlaIleSerLysAlaProIleGluSerTyrArgMet 87  
QY 2373 GTGAATAAGGAAACGCTGTTGAAGCGGCGGCGGCGACGATCTGAATATCGGCACA 2432  
Db 88 ValAsnLysGluThrLeuLeuGluGlyAlaLysArgAlaHisAspLeuAsnIleGlyThr 107  
QY 2433 TATTGTATCGTGGCAAGCGGACAGAGTCCGCTTAACAGAGAAGTGGATCGGTAGAT 2492  
Db 108 TyrCysIleValAlaSerGlyArgGlyProSerAsnArgGluValAspGlnValValAsp 127  
QY 2493 GCGGTTACGAGAAATTAAGAGACGATGCGTCTTGAAGTTCGATGCTTTGGACTGTG 2552  
Db 128 AlaValGlnGluIleLysGluThrTyrGlyLeuLysValCysAlaCysLeuGlyLeuLeu 147  
QY 2553 AAGCCAGACGAGCGGCGGCTCAAGAGTACGAGGAGTACCGCTATATCAATAATTG 2612  
Db 148 LysProGlnGlnAlaLysArgLeuLysAspAlaGlyValAspArgTyrAsnHisAsnLeu 167  
QY 2613 AATACGTCACAGAGAAACCATTTCAACATCACACCTCACATACATACAGTACAGAGTC 2672  
Db 168 AsnThrSerGlnArgAsnHisSerAsnIleThrThrSerHisThrTyrAspAspArgVal 187  
QY 2673 AATACGTTGAATCGCAAAAGAAATCGGGGCTGTCTCGGCTTACCGGCGCCATTCGGG 2732  
Db 188 AsnThrValGluIleAlaLysGluSerGlyLeuSerProCysSerGlyAlaIleIleGly 207

```
Qy 2733 ATGAGGAGCGAAGACAGGATGTCTATTCACATCGCCAAAAGCTTGAAGGCTCTTGACGCG 2732
Db 208 IleybGluThrLysGlnAspValIleAspIleAlaLysSerLeuLysAlaLeuAspAla 227
Qy 2793 GATTCACATTCCTGTAATTTTTTTCATGCAATTCATGTCACGCGCTTAGAAGGCGTCAAC 2852
Db 228 AspSerIleProValAsnPheLeuHisAlaIleAspGlyThrProLeuGluGlyValAsn 247
Qy 2853 GAATTAACCCGCTGTATTTGTTTAAAGTGTGGCGTGTTCGGTTTATCATCCATCA 2912
Db 248 GluLeuAsnProLeuTyrCysLeuLysValLeuAlaLeuPheArgPheIleAsnProSer 267
Qy 2913 AAAGAAATTCGATTTCCGGAGGAGAGAGTCAATCTCCGACATTCGACGCCATTAGGG 2972
Db 268 LysGluIleArgIleSerGlyGlyArgGluValAsnLeuArgSerLeuGlnProLeuGly 287
Qy 2973 CTTTACGCGCGCAAACTCCATTTTTTGTCTCGAGACTACTTTAACTGCGCGGCAAGAGGAG 3032
Db 288 LeuTyrAlaAlaAsnSerIlePheValGlyAspTyrLeuThrAlaGlyGlnGluGlu 307
Qy 3033 ACGGAGGATCATAAATCTGAGTGATTTAGGCTTTGAAGTTGAATCATGTCGAAGAATG 3092
Db 308 ThrGluAspHisLysMetLeuSerAspLeuGlyPheGluValGluSerValGluGluMet 327
Qy 3093 AAGCGTAGTTTAAGTCGGAAGC 3116
Db 328 LysAlaSerLeuSerAlaLysSer 335

RESULT 13
Q65ML1_BACLD
ID Q65ML1_BACLD PRELIMINARY; PRT; 379 AA.
AC Q65ML1_Q62Y04;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 13-SEP-2005 (TREMBLrel. 31, Last annotation update)
DE BioF (8-amino-7-oxononanoate synthase).
GN Name:bioF; OrderedLocusNames=BL00954; BL00768;
OS Bacillus licheniformis (strain DSM 13 / ATCC 14580).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=279010;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15383718; DOI=10.1159/000079829;
RA Veith B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
RA Ehrenreich P., Baumer S., Henne A., Liesegang H., Merkl R.,
RA Ehrenreich A., Gottschalk G.;
RT "The complete genome sequence of Bacillus licheniformis DSM13, an
RT organism with great industrial potential.";
RL J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15461803; DOI=10.1186/gb-2004-5-10-r77;
RA Rey M.W., Ramaiya P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J.,
RA Tang M., Lopez de Leon A., Xiang H., Gueti V., Clausen I.G.,
RA Olsen P.B., Rasmussen M.D., Andersen J.T., Joergensen P.L.,
RA Larsen T.S., Sorokin A., Bolotin A., Lapidus A., Galleron N.,
RA Ehrlich S.D., Berka R.M.;
RT "Complete genome sequence of the industrial bacterium Bacillus
RT licheniformis and comparisons with closely related Bacillus species.";
RL Genome Biol. 5:RESEARCH077.1-RESEARCH077.12(2004).
DR EMBL; AE017333; AAU39703.1; -; Genomic DNA.
DR EMBL; CP000002; AAU22354.1; -; Genomic DNA.
DR GO; GO:0008710; F:8-amino-7-oxononanoate synthase activity; IEA.
DR GO; GO:0016769; P:transferase activity, transferring nitrogen. .; IEA.
DR GO; GO:0009058; P:biotin synthesis; IEA.
DR GO; GO:0009102; P:biotin biosynthesis; IEA.
DR InterPro; IPR004839; AminoTrans_I/II.
DR InterPro; IPR001917; AminoTrans_II.
DR InterPro; IPR004723; BioF.
DR Pfam; PF00155; AminoTrans_1_2; 1.
DR TIGRFAMs; TIGR00858; bioF; 1.
DR PROSITE; PS00599; AA_TRANSF_CLASS_2; UNKNOWN_1.
```

```
KW Complete proteome.
SQ SEQUENCE 379 AA; 41209 MW; C9CB203EA08F20AF CRC64;

Alignment Scores:
Pred. No.: 4,07e-89 Length: 379
Score: 1286.50 Matches: 257
Percent Similarity: 82.8% Conservative: 41
Best Local Similarity: 71.4% Mismatches: 51
Query Match: 23.3% Indels: 11
DB: 2 Gaps: 2

US-10-681-086-1 (1-3156) x Q65ML1_BACLD (1-379)
Qy 22 TTGAAGATTGATCTCTGTTTAAAGAGCGGTAGACAGAATGAAGAAGCGCGGTACAT 81
Db 1 MetArgIleAspGluTrpLeuSerArgLeuAlaLysThrLysAlaIleGlyLeuTyr 20
Qy 82 CGTAACCTCGCGTCAATGTGATGGAGCCCGCTTCCAGAGAGGATATTTGATGGCGAAAT 141
Db 21 ArgSerLeuLys-----LeuProGlnAlaGluArg----- 30
Qy 142 CAACGGCTGTGCTCTCAACAATTTATTAGGCTCCGAAGCGATAGACGTTTTCATCAT 201
Db 31 ---ThrAsnTrpAlaSerAsnAspTyrLeuGlyLeuAlaAsnAspLysArgLeuIleHis 49
Qy 202 GCAGCCCAACACGATTCGACCAATTTGGGACAGAGCAGCGTTCCACGTTTAAACGACA 261
Db 50 AlaAlaGluThrAlaLeuArgArgPheGlyAlaGlySerThrGlySerArgLeuThrSer 69
Qy 262 GGCAATTCGGTCTCGCATGAAAGAGCTAGAAAAGAGATTGCCAGCTTTTAAACTGACAGA 321
Db 70 GlyAsnThrAlaTrpHisGluLysLeuGluArgLysIleAlaGlyPheLysGlnThrGlu 89
Qy 322 GCGGCCCTGCTGTTTTCGACGGTTACTTGGCCCAATGTCGTTGTCCTTTCATCTTCGCCA 381
Db 90 AlaAlaLeuLeuPheSerSerGlyTyrLeuAlaAsnValGlyValLeuSerSerLeuPro 109
Qy 382 GAAAGGAAGATGTCTATTAAAGTACGACGCTCAATCATGCAAGTATGATCGACGGCTGC 441
Db 110 GluLysGlyAspValIleLeuSerAspGlnLeuAsnHisAlaSerIleIleAspGlyCys 129
Qy 442 CGACTTTTAAAGGCTGATCAGATTGTTTATCGGCATATTGATATGAATGATCTTGAANAAC 501
Db 130 ArgLeuSerLysAlaAspThrValValTyrArgHisLysAspMetAsnAspLeuGluGlu 149
Qy 502 AAGCTGAATGAACACACAGCGTTTATCAGCGCGCTTTTATCGTAACAGACGGAGTATTCAGC 561
Db 150 LysLeuArgAlaAlaGlnSerArgAlaArgPheIleValThrAspGlyValPheSer 169
Qy 562 ATGGATGGCAATTCGCGCCCTCTTGTATCAGATCATCTCCTTGGGAAACGCTATCATGCC 621
Db 170 MetAspGlyThrIleAlaProLeuAspGluIleLeuLeuAlaLysGlnTyrArgAla 189
Qy 622 TTCGTGTCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 681
Db 190 PheValIleValAspAspAlaHisAlaThrGlyValLeuGlyGluAlaGlyArgGlyThr 209
Qy 682 AGTGAATACCTTGGTGTGTTGTCGCGACATTTGTTATCGGCACCTTAAAGCAAGCTGTGGC 741
Db 210 SerGluTyrPheGlyValSerProAspValValIleGlyThrLeuSerLysAlaValGly 229
Qy 742 GCGGAAGGAGGTTTGGCGGAGGATCAGCGTCTTTCATCGACTTTTGTGCAACCATGCC 801
Db 230 AlaGluGlyGlyPheValAlaGlySerLysThrLeuIleAspPheLeuLeuAsnHisAla 249
Qy 802 AGAACATTATCTTTCAACCGCTATTTCGCCAGCCAGCTGTGCGGTGCTTCACAGAGCT 861
Db 250 ArgThrPheIlePheGlnThrAlaValProAlaSerCysAlaAlaAlaCysLysAla 269
Qy 862 TTCACATCATTTGAGCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 921
Db 270 LeuAspIleIleGluAspSerArgAlaLysArgArgLeuLeuGlnSerSerValAsnThr 289
```

```
QY 922 ATCAGAACCACTGTAAGAATATGGTATATGTGTGAAGAGAGATACACACCATATT 981
Db 290 IleyysargSerLeuValAspIleGlyPheThrValaAenGlyGluAspThrProIleile 309
QY 982 CCTAGTCATGCGGATGCCATAAACGGTCTCTATTGCTGAAACAACTGCAGGGCAAG 1041
Db 310 ProValMetIleGlyAspProGlnLysAlaValGlnPheAlaAenGlyLeuLysGluLys 329
QY 1042 GGAATTTATGCTCTGCGCATTCGGCGCCCAACCGTTGCGCGGGTGAAGCCGGATTGCA 1101
Db 330 GlyIleGluAlaProAlaIleArgProThrValAlaGluGlyGluSerArgIleArg 349

RESULT 14
Q75WE9_922ZZ PRELIMINARY; PRT; 253 AA.
AC Q75WE9;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Highly thermostable kanamycin nucleotidyltransferase
DE (EC 2.7.7.46).
GN Name=htk;
OS Plasmid PUB110.
OC other sequences; plasmids.
OX NCBI_TaxID=2655;
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=20013005; PubMed=10544290;
RA Haseki J., Yano T., Koyama Y., Kuramitsu S., Kagamiyama H.;
RT "Directed evolution of thermostable kanamycin-resistance gene: a
RT convenient selection marker for Thermus thermophilus.";
RL J. Biochem. 126:951-956(1999).
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=21483306; PubMed=11602251; DOI=10.1016/S0014-5793(01)02926-X;
RA Hashimoto Y., Yano T., Kuramitsu S., Kagamiyama H.;
RT "Disruption of Thermus thermophilus genes by homologous recombination
RT using a thermostable kanamycin-resistant marker.";
RL FEBS Lett. 506:231-234(2001).
RN NUCLEOTIDE SEQUENCE.
RA Haseki J., Yano T., Koyama Y., Kuramitsu S., Kagamiyama H.;
RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB121443; BAC98411.1; -; Genomic_DNA.
DR SMR; Q75WE9; 1-253.
DR GO; GO:0016779; F:nucleotidyltransferase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR002934; DNA_pol_beta.
DR Pfam; PF01909; NTP_transf_2; 1.
KW Nucleotidyltransferase; Transferase.
SQ SEQUENCE 253 AA; 28860 MW; BC691E32B8E5705E CRC64;

Alignment Scores:
Pred. NO.: 3.1e-84 Length: 253
Score: 1222.00 Matches: 231
Percent Similarity: 93.7% Conservative: 6
Best Local Similarity: 91.3% Mismatches: 16
Query Match: 22.1% Indels: 0
DB: Gaps: 0

US-10-681-086-1 (1-3156) x Q75WE9_922ZZ (1-253)

QY 1579 GTGAATGACCAATAAATGACTAGAGAGAAAGATGTTTCATGAATTAAG 1638
Db 1 MetLysGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluLys 20
QY 1639 GAACGAATATTCGATAAATATGGGATGATGTTAAGGCTATTGGTGTATTGGCTCTCT 1698
Db 21 GluArgIleLeuAspLysTyArgAspValLysAlaIleGlyValTyGlySerLeu 40
QY 1699 GGTGCTGAGATGATGGGCCCTATTTCGATATTCGATGATGTCATGTCACAGAG 1758
Db 41 GlyArgGlnThrAspGlyProTySerAspIleGluMetMetCysValLeuSerThrGlu 60
```

```
QY 1759 GAAGCAGAGTTCCAGCCATGAATGAGCAACCGTGAGTGGAGGTGAAGTGAATTTTGAT 1818
Db 61 GlyValGluPheSerTyGluTrpThrThrGlyGluTrpLysAlaGluValAsnPheTy 80
QY 1819 AGCGAAGAGATTTCTACTAGATTATGATCTCAGGTGAATCAGATTGGCCCGCTTACACAT 1878
Db 81 SerGluGluIleLeuLeuAspTyArgSerArgValGluProAspTrpProLeuThrHis 100
QY 1879 GGTCAATTTTCTCTATTGTCGGATTTATGATTTCAGGTGGATATCTTAGAGAAAGTGTAT 1938
Db 101 GlyArgPhePheSerIleLeuProIleTyAspProGlyGlyTyPheGluLysValTy 120
QY 1939 CAAACTCTAAATCGGTAGAACCCCAACGTTCCACGATCGATTGTCCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140
QY 1999 GAAGAGCTGTTGATATGATGAGCAAAATGCGTAATATTGCTGTGCAAGGCCACACACA 2058
Db 141 GluGluLeuPheGluTyAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
QY 2059 TTTCTACCATCTTGACTGTACAGTAGCAATGCGAGTGCATGTTGATTGGTCTGCAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
QY 2119 CATCGCATCTGTTATACGACGAGCGCTTCGCTCTTAACCTGAAGCAGTTAAGCAATCAGAT 2178
Db 181 HisArgIleCysTyThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp 200
QY 2179 CTTCTTCAGGTTATGACCATCTGTGCGCAGTTCGTATGCTGTCACTTCGACTCT 2238
Db 201 LeuProGlyTyTrpValGlnLeuCysGlnLeuValMetSerGlyGlnLeuSerAspPro 220
QY 2239 GAGAACTCTTCGAATCGTAGAGAAATTCCTGGAATGGATTTCAGGAGTGCACAGAACGA 2298
Db 221 GluLysLeuLeuGluSerLeuLysPheTrpAsnGlyValGlnGluTrpAlaGluArg 240
QY 2299 CACGATATATAGTGGATGTGTCAAACGCAATACCATTT 2337
Db 241 HisGlyTyTrpIleValAspValSerLysArgIleProPhe 253

RESULT 15
Q65MK9_BACLD
ID Q65MK9_BACLD PRELIMINARY; PRT; 333 AA.
AC Q65MK9; Q62Y02;
DT 25-OCT-2004 (TReMBLrel. 28, Created)
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DT 13-SEP-2005 (TReMBLrel. 31, Last annotation update)
DE BioB (Biotin synthetase).
GN Name=bioB; OrderedLocusNames=BL00956, BLi00770;
OS Bacillus licheniformis (strain DSM 13 / ATCC 14580).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=279010;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15383718; DOI=10.1159/000079829;
RA Veith B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
RA Ehrenreich P., Baeumer S., Henne A., Liesegang H., Merkl R.,
RA Ehrenreich A., Gottschalk G.;
RT "The complete genome sequence of Bacillus licheniformis DSM13, an
RT organism with great industrial potential.";
RL J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RX PubMed=15463803; DOI=10.1186/gb-2004-5-10-r77;
RA Rey M.W., Ramaiya P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J.,
RA Tang M., Lopez de Leon A., Xiang H., Gusti V., Clausen I.G.,
RA Olsen P.B., Rasmussen M.D., Andersen J.T., Joergensen P.L.,
RA Larsen T.S., Sorokin A., Bolotin A., Lapidus A., Galleron N.,
RA Ehrlich S.D., Berk R.M.;
RT "Complete genome sequence of the industrial bacterium Bacillus
RT licheniformis and comparisons with closely related Bacillus species.";
RL Genome Biol. 5:RESEARCH077.1-RESEARCH077.12(2004).
```





GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: February 10, 2006, 02:56:51 ; Search time 19 Seconds  
(without alignments)

2746.573 Million cell updates/sec

Title: US-10-681-086-1

Perfect score: 5517

Sequence: 1 ggatccacgaggttacgagc.....atcggtatgatgcgaattc 3156

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Xgapop 10.0 , Xgapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 1144120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp

-Q=/abs/ARSSWEB\_spool/US10681086/runat\_09022006\_140520\_11815/app\_query.fasta\_1

-DB-issued Patents AA -Qm=fastan -SUFFIX=n2p.rai -MINMATCH=0.1 -LOOPCL=0

-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human4.0.cdi

-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15

-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000

-HOST=abs02p -USPR=US10681086 -CGN 1 1 71 @runat\_09022006\_140520\_11815

-NCPU=6 -ICPU=3 -NO\_MMAP -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DBLOP=6 -DELEXT=7

Database : Issued Patents AA:

1: /cgn2\_6/ptodata/1/iaa/5 COMB.pap.\*

2: /cgn2\_6/ptodata/1/iaa/6 COMB.pap.\*

3: /cgn2\_6/ptodata/1/iaa/H COMB.pap.\*

4: /cgn2\_6/ptodata/1/iaa/PCITUS COMB.pap.\*

5: /cgn2\_6/ptodata/1/iaa/RE COMB.pap.\*

6: /cgn2\_6/ptodata/1/iaa/backfiles1.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1333	24.2	253	1	US-08-642-045B-4
2	1333	24.2	253	2	US-08-852-268-2
3	1333	24.2	253	2	US-03-697-186B-11
4	1318	23.9	253	2	US-03-697-186B-1
5	1282	23.2	253	2	US-03-697-186B-12
6	1277	23.1	253	2	US-03-697-186B-14
7	1268	23.0	253	2	US-03-697-186B-15
8	1267	23.0	253	2	US-03-697-186B-2
9	1267	23.0	253	2	US-03-697-186B-13
10	1265	22.9	253	2	US-03-697-186B-20
11	1262	22.9	253	2	US-03-697-186B-17
12	1262	22.9	253	2	US-03-697-186B-18

13	1260	22.8	253	2	US-09-697-186B-16
14	1256	22.8	253	2	US-09-697-186B-19
15	1222	22.1	253	2	US-09-697-186B-3
16	942	17.1	387	2	US-08-935-263-6
17	942	17.1	387	2	US-09-594-185-6
18	942	17.1	387	2	US-10-033-078-6
19	942	17.1	387	2	US-10-763-933-6
20	851	15.4	338	2	US-08-935-263-8
21	851	15.4	338	2	US-09-594-185-8
22	851	15.4	338	2	US-10-033-078-8
23	851	15.4	338	2	US-10-763-933-8
24	828	15.0	327	2	US-09-134-001C-3471
25	828	15.0	327	2	US-09-710-279-540
26	828	15.0	332	2	US-09-710-279-2562
27	704	12.8	414	2	US-09-107-532A-5771
28	673.5	12.2	398	2	US-08-935-263-12
29	673.5	12.2	398	2	US-09-594-185-12
30	673.5	12.2	398	2	US-10-033-078-12
31	673.5	12.2	398	2	US-10-763-933-12
32	660	12.0	512	2	US-09-902-540-13924
33	612.5	11.1	488	2	US-09-902-540-11131
34	602.5	10.9	387	2	US-09-180-109A-2
35	602.5	10.9	387	2	US-09-180-109A-6
36	602.5	10.9	387	2	US-09-180-109A-8
37	591.5	10.7	412	2	US-09-489-039A-10771
38	590.5	10.7	436	2	US-09-252-991A-19792
39	587.5	10.6	411	2	US-09-543-681A-7404
40	567.5	10.3	771	2	US-09-712-363-157
41	559	10.1	391	2	US-09-543-681A-5575
42	546.5	9.9	394	2	US-09-328-352-4736
43	535	9.7	384	1	US-08-401-068-2
44	535	9.7	384	1	US-08-846-338-2
45	535	9.7	384	2	US-08-411-768B-7

#### ALIGNMENTS

#### RESULT 1

US-08-642-045B-4

; Sequence 4, Application US/08642045B

; Patent No. 5851804

; GENERAL INFORMATION:

; APPLICANT: Snyder, Linda A.

; APPLICANT: Satishchandra, C.

; TITLE OF INVENTION: CHIMERIC KANAMYCIN RESISTANCE GENE

; NUMBER OF SEQUENCES: 25

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5851804ris

; STREET: One Liberty Place, 46th floor

; CITY: Philadelphia

; STATE: Pennsylvania

; COUNTRY: USA

; ZIP: 19103

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: WINDOWS

; SOFTWARE: WordPerfect 6.0/6.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/642,045B

; FILING DATE: 06-MAY-1996

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: DeLuca, Mark

; REGISTRATION NUMBER: 33,229

; REFERENCE/DOCKET NUMBER: APOL-0262

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 215-568-3100

; TELEFAX: 215-568-3439

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 253 amino acids

; TYPE: amino acid

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-642-045B-4

Alignment Scores:
Pred. No.: 6.5e-138 Length: 253
Score: 1333.00 Matches: 252
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.6% Mismatches: 0
Query Match: 24.2% Indels: 0
DB: 1 Gaps: 0

US-10-681-086-1 (1-3156) x US-08-642-045B-4 (1-253)

QY 1579 GTGAATGACCAATAATAATGACTAGAGAGAAAGAAATGAAGATTGTTTCATGAATTAAG 1638
Db 1 MetAenGlyProIleMetThrArgGluArgMetLysIleValHisGluIleLys 20

QY 1639 GAACGAATATTGGATAAATATGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCTT 1698
Db 21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40

QY 1699 GGTCTCAGACTGATGGCCCTATTTCGGATATTGAGATGATGTGTTCATGTCAACAGAG 1758
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60

QY 1759 GAAGCAGAGATTCTACTAGATTATGCATCTCAGGTGGAATCAGATTGCCCGCTTACACAT 1818
Db 61 GluAlaGluPheSerHisGluIleThrThrGlyGluTrpLysValGluValAsnPheAsp 80

QY 1819 AGCGAAGAGATTCTACTAGATTATGCATCTCAGGTGGAATCAGATTGCCCGCTTACACAT 1878
Db 81 SerGluGluIleLeuAspTyrAlaSerGlnValGluSerAspTrpProLeuThrHis 100

QY 1879 GGTCAATTTTCTCTATTTCGCCGATTTATGATTTCAGTGGATACCTTAGAGAAAGTGTAT 1938
Db 101 GlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120

QY 1939 CAAAGCTGAATCGGTAGACCCCAACGCTCCACGATGGATTGTGCCCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAlaLeuIleVal 140

QY 1999 GAAGAGCTGTTTGAATATCGAGCAAAATGGCGTAAATATTTCGTGTCAAGGACCCGACACA 2058
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160

QY 2059 TTTCTACCATCTTGCATGTACAGTAGCAATGGCAGGTGCCATGTTGATTGGTCTGCAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180

QY 2119 CATCGCATCTGTATACGACGAGCGCTTCGGTCTTAAGTGAAGCAGTTAAGCAATCAGAT 2178
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200

QY 2179 CTTCTCCTCAGGTATGACCATCTGTGCCAGATTGCTGTAATGCTGTGCAACTTTCGACCT 2238
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220

QY 2239 GAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGAGTGGACGAGAACGA 2298
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240

QY 2299 CACGGATATAGTGGATGTCTCAAAACGCATACCATTT 2337
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 2
US-08-852-268-2
; Sequence 2, Application US/08852268
; Patent No. 6143527
; GENERAL INFORMATION:
; APPLICANT: Pachuk, Catherine J.
; APPLICANT: Samuel, Mano
```

```
QY 1939 CAAACTGCTAAATCGGTAGAGCCCAACGCTTCACGATGCGATTTGTGCCCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAlaLeuIleVal 140
QY 1999 GAAGAGCTGTTGAATATGACGAGCAATGCGGTAATATTCGTGTGCAAGGACCGACAACA 2058
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
QY 2059 TTTCTACCATCTTGTACGCTGACAGGTAGCAATGCGAGTGCCTATGTTGATTGGTCTGCAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
QY 2119 CATCGCATCTGTTATACGACGAGCGCTTCGGTCTTAATCAAGCAGTAAAGCAATCAGAT 2178
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200
QY 2179 CTTCTCCTCAGGTTATGACCATCTGTGCGCAGTTCGTAATGCTGGTCAACTTTCGCACTCT 2238
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
QY 2239 GAGAACTCTCGAATCGCTAGAGAAATTCCTGGAATGGGATTCAGGAGTGCAGACGACGA 2298
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
QY 2299 CACGGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 3
US-09-697-186B-11
; Sequence 11, Application US/09697186B
; Patent No. 6723543
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, SIGEYUKI
; APPLICANT: HOSEKI, JUN
; APPLICANT: YANO, TAKATO
; APPLICANT: KOYAMA, YOSHINORI
; APPLICANT: KURAMITSU, SEIKI
; APPLICANT: KAGAMIYAMA, HIROYUKI
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
; FILE REFERENCE: 04853.0048-00000
; CURRENT FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US/09/697,186B
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-697-186B-11

Alignment Scores:
Pred. No.: 6.5e-138 Length: 253
Score: 1333.00 Matches: 252
Percent Similarity: 100.0% Conservatives: 1
Best Local Similarity: 99.6% Mismatches: 0
Query Match: 24.2% Indels: 0
DB: 2 Gaps: 0

US-10-681-086-1 (1-3156) x US-09-697-186B-11 (1-253)

QY 1579 GTGAATGACCAATAATAAGTACGTAGAGAAAGATGAAGATTGTTCAATGAATTAAG 1638
Db 1 MetAsnGlyProIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
QY 1639 GAACGAATATGGTAATAATGCGGATCATGTTAAGCTATTGGTGTATTGGCTCCTT 1698
Db 21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40
QY 1699 GGTGCTCAGACTGATGGCCCTATTTCGGATATTGAGATGATGTGTGTCATGTCAACAGAG 1758
```

```
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60
QY 1759 GAAGCAGAGTTCCACCATGATCGCAACCGGTGAGTGGAGGTGGAAGTCAATTTTGAT 1818
Db 61 GluAlaGluPheSerHisGluTrpThrThrGlyLysValGluValAsnPheAsp 80
QY 1819 AGCGAAGAGATCTTACTAGATTATGCATCTCAGGTGGAATCAGATTGGCCGCTTACACAT 1878
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpProLeuThrHis 100
QY 1879 GGTCAATTTTCTCTATTTTGGCGATTATGATTCAGGTGGATCATTAGAGAAAGTGTAT 1938
Db 101 GlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120
QY 1939 CAACTGCTAAATCGGTAGAGCCCAACGTTCCAGTCAGATGCGATTTGTGCCCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAlaLeuIleVal 140
QY 1999 GAAGAGCTGTTTGAATATGCGAATAATGCGGTAATATTCGTGTGCAAGGACCGACAACA 2058
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
QY 2059 TTTCTACCATCTTGTACGCTGACAGGTAGCAATGCGAGTGCCTATGTTGATTGGTCTGCAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
QY 2119 CATCGCATCTGTTATACGACGAGCGCTTCGGTCTTAATCAAGCAGTAAAGCAATCAGAT 2178
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200
QY 2179 CTTCTCCTCAGGTTATGACCATCTGTGCGCAGTTCGTAATGCTGGTCAACTTTCGCACTCT 2238
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
QY 2239 GAGAACTCTCTGGAATCGCTAGAGAAATTCCTGGAATGGGATTCAGGAGTGCAGACGACGA 2298
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
QY 2299 CACGGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 4
US-09-697-186B-1
; Sequence 1, Application US/09697186B
; Patent No. 6723543
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, SIGEYUKI
; APPLICANT: HOSEKI, JUN
; APPLICANT: YANO, TAKATO
; APPLICANT: KOYAMA, YOSHINORI
; APPLICANT: KURAMITSU, SEIKI
; APPLICANT: KAGAMIYAMA, HIROYUKI
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
; FILE REFERENCE: 04853.0048-00000
; CURRENT FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US/09/697,186B
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme
; OTHER INFORMATION: obtained by introduction of point mutation into
; OTHER INFORMATION: wild type KNT gene of Staphylococcus aureus and
; OTHER INFORMATION: its expression
US-09-697-186B-1
```

## Alignment Scores:

Pred. No.: 2,96e-136 Length: 253  
Score: 1318.00 Matches: 250  
Percent Similarity: 99.2% Conservative: 1  
Best Local Similarity: 98.8% Mismatches: 2  
Query Match: 23.9% Indels: 0  
DB: Gaps: 0

US-10-681-086-1 (1-3156) x US-09-697-186B-1 (1-253)

QY 1579 GTGAATGACCAATAATATGACTAGACAGAAAGAAATGAAGATTGTTTCATGAATTAAG 1638  
Db 1 MetAenGlyProIleMetThrArgGluGluArgMetLysAlaValHisGluLeuLys 20  
QY 1639 GAACCAATATGCAATAATATGGGATGATGTTAAGGCTATTTGGTTTATGGCTCTCT 1698  
Db 21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40  
QY 1699 GGTCTGTACAGCTGATGGGCCCTATTTCGGATATTGAGATGATGTGTCTCATGTCAACAGAG 1758  
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetCysValMetSerThrGlu 60  
QY 1759 GAACGAGAGTTTCAGCCATGAATGGACAAACCGGTGAGTGGAGGTGGAAAGTGAATTTTGTAT 1818  
Db 61 GluAlaGluPheSerHisGluTyrThrGlyGluTyrLysValGluValAsnPheTyr 80  
QY 1819 AGCGNAGAGATCTACTAGATATGCACTCAGGTGGAATCAGATTGGCCCTTACACAT 1878  
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTyrProLeuThrHis 100  
QY 1879 GGTCAATTTTCTCTATTTCGCGATTTATGATTCAGGTGATATCTAGAGAAAGTGTAT 1938  
Db 101 GlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120  
QY 1939 CAAACTGCTAAATCGGTAGAACCCAAACGTTTCCACGATGCGATTTGTGCCCTTATCGTA 1998  
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140  
QY 1999 GAAGAGCTTTGATATGACGCAATGCGTAATATTCGTGTCAGAGGACCCAGACACA 2058  
Db 141 GluGluLeuPheGluTyrAlaGlyLysTyrArgAsnIleArgValGlnGlyProThrThr 160  
QY 2059 TTTCTACCATCTCTGACTGTACAGGTAGCAATGGCAGGTGCGCATGTTGATTTGGTCTGCAT 2118  
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180  
QY 2119 CATCGCATCTGTTATACGACGAGCGCTTCGTTTAACTGAAGCAGTTTAAAGCAATTCAGAT 2178  
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200  
QY 2179 CTTCCCTTCAGGTATGACCATCTGCGCAGTTCGTAATGTCGTCAACTTTCGACTCT 2238  
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220  
QY 2239 GAGAACTTCCTGGAATCGCTAGAGAAATTTCTGGAATGGATTTCAGAGGTGACAGAGACA 2298  
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTyrAsnGlyIleGlnGlnTyrThrGluArg 240  
QY 2299 CACGGATATAGTGGATGTGTCAAAACGCGATACCATTT 2337  
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

## RESULT 5

US-09-697-186B-12  
; Sequence 12, Application US/09697186B  
; Patent No. 6723543  
; GENERAL INFORMATION:  
; APPLICANT: YOKOYAMA, SIGEYUKI  
; APPLICANT: HOSEKI, JUN  
; APPLICANT: YANO, TAKATO  
; APPLICANT: KODAMA, YOSHINORI  
; APPLICANT: KURAMITSU, SEIKI

; APPLICANT: KAGAMIYAMA, HIROYUKI  
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD  
; FILE REFERENCE: OF SCREENING THERMOPHILIC BACTERIA USING THE SAME  
; CURRENT APPLICATION NUMBER: US/09/697,186B  
; CURRENT FILING DATE: 2000-10-27  
; PRIOR APPLICATION NUMBER: JP 309616/1999  
; PRIOR FILING DATE: 1999-10-29  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme  
; OTHER INFORMATION: obtained by introduction of point mutation into  
; OTHER INFORMATION: wild type KNT gene of Staphylococcus aureus and  
; OTHER INFORMATION: its expression

US-09-697-186B-12

## Alignment Scores:

Pred. No.: 2,82e-132 Length: 253  
Score: 1282.00 Matches: 241  
Percent Similarity: 97.6% Conservative: 6  
Best Local Similarity: 95.3% Mismatches: 6  
Query Match: 23.2% Indels: 0  
DB: Gaps: 0

US-10-681-086-1 (1-3156) x US-09-697-186B-12 (1-253)

QY 1579 GTGAATGACCAATAATATGACTAGACAGAAAGAAATGAAGATTGTTTCATGAATTAAG 1638  
Db 1 MetAenGlyProIleMetThrArgGluGluArgMetLysAlaIleValTyrGluLeuLys 20  
QY 1639 GAACCAATATGGAATAAATATGGGATGATGTTAAGGCTATTTGTGTTTATGGCTCTCTT 1698  
Db 21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40  
QY 1699 GGTCTGTACAGCTGATGGGCCCTATTTCGGATATTGAGATGATGTGTGTCATGTCACAGAG 1758  
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetCysValMetSerThrGlu 60  
QY 1759 GAACGAGAGTTTCAGCCATGAATGGACAAACCGGTGAGTGGAGGTGGAAGTCAATTTTTCAT 1818  
Db 61 GluAlaGluPheSerHisGluTyrThrGlyGluTyrLysAlaGluValAsnPheTyr 80  
QY 1819 AGCGNAGAGATCTACTAGATATTGCAATCTCAGGTGGAATCAGATTGGCCCTTACACAT 1878  
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluSerAspTyrProLeuThrHis 100  
QY 1879 GGTCAATTTTCTCTATTTCGCGATTTATGATTCAGGTGATATCTAGAGAAAGTGTAT 1938  
Db 101 GlyArgPhePheSerIleLeuProIleTyrAspProGlyGlyTyrPheGluLysValTyr 120  
QY 1939 CAAACTGCTAAATCGGTAGAACCCAAACGTTTCCACGATGCGATTTGTGCCCTTATCGTA 1998  
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140  
QY 1999 GAAGAGCTTTGATATGACGCAATGCGGTAATATTCGTGTGCAAGGACCCAGACACA 2058  
Db 141 GluGluLeuPheGluTyrAlaGlyLysTyrArgAsnIleArgValGlnGlyProThrThr 160  
QY 2059 TTTCTACCATCTCTGACTGTACAGGTAGCAATGGCAGGTGCGCATGTTGATTTGGTCTGCAT 2118  
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180  
QY 2119 CATCGCATCTGTTATACGACGAGCGCTTCGTTTAACTGAAGCAGTTTAAAGCAATTCAGAT 2178  
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaLeuLysGlnSerAsp 200  
QY 2179 CTTCCCTTCAGGTATGACCATCTGTCGAGTTCGTAATGTCGTCAACTTTCGACTCT 2238  
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220  
Qy 2239 GAGAACTCTCGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGCACAGACGA 2298  
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyValGlnGluTrpAlaGluArg 240  
Qy 2299 CACGATATATAGTGGATGTCTAAAACGCATACATTT 2337  
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 6  
US-09-697-186B-14  
; Sequence 14, Application US/09697186B  
; Patent No. 6723543  
; GENERAL INFORMATION:  
; APPLICANT: YOKOYAMA, SIGEYUKI  
; APPLICANT: HOSEKI, JUN  
; APPLICANT: YANO, TAKATO  
; APPLICANT: KOYAMA, YOSHINORI  
; APPLICANT: KURAMITSU, SEIKI  
; APPLICANT: KAGAMIYAMA, HIROYUKI  
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD  
; TITLE OF INVENTION: OF SCREENING THERMOPHILIC BACTERIA USING THE SAME  
; FILE REFERENCE: 04853.0048-00000  
; CURRENT APPLICATION NUMBER: US/09/697,186B  
; PRIOR FILING DATE: 2000-10-27  
; PRIOR FILING DATE: 1999-10-29  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 14  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme  
; OTHER INFORMATION: obtained by introduction of point mutation into  
; OTHER INFORMATION: wild type KNT gene of Staphylococcus aureus and  
; OTHER INFORMATION: its expression  
US-09-697-186B-14

Alignment Scores:  
Pred. No.: 1.01e-131 Length: 253  
Score: 1277.00 Matches: 241  
Percent Similarity: 97.6% Conservative: 6  
Best Local Similarity: 95.3% Mismatches: 6  
Query Match: 23.1% Indels: 0  
DB: 2 Gaps: 0

US-10-681-086-1 (1-3156) x US-09-697-186B-14 (1-253)

Qy 1579 GTGAATGGACCAATAATACTAGACAGAGAAAGATGAGATTGTCATGAAATTAAG 1638  
Db 1 MetAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20  
Qy 1639 GAACGAATATGGATAAATATGGGATCATGTTAAGGCTATTTGGTTTATGGCTCTCTT 1698  
Db 21 GluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyrGlySerLeu 40  
Qy 1699 GGTGCTCAGACTGATGGCCCTATTCCGATATTGAGATGATGTGTCTCATGTCACAGAG 1758  
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60  
Qy 1759 GAACGACGATTCAGCCATGAATGACACCGGTGAGTGGAGGTGGAAGTCAATTTTGTAT 1818  
Db 61 GlyAlaGluPheSerTyrGluTrpThrThrGlyGluTrpLysAlaGluValAsnPheTyr 80  
Qy 1819 AGCGAAGAGATTCCTACTAGATTATGATCTCAGGTGGAATCAGATTGGCGCTTACACAT 1878  
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluProAspTrpProLeuThrHis 100  
Qy 1879 GGTCAATTTTCTCTATTGTCGCAATTTATGATTCAGGTGATCTTACAGAAAGTGTAT 1938  
Db 1938

Db 101 GlyArgPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120  
Qy 1939 CAAACTCTAAATCGGTAGAGCCAAACGTTCCACGATCGATTGTCCTTATCGTA 1998  
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140  
Qy 1999 GAAGAGCTGTTTGAATATGACAGCAAAATGGCGTAATATTCTGTGTCAAGGACCGACACA 2058  
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160  
Qy 2059 TTTCTACCATCTGACTGACAGTAGCAATGGCAGATGCCAGTCCCATGTTGATGGTCTCGAT 2118  
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180  
Qy 2119 CATCGCATCTGTTATACGACGAGCGCTTCGTCTTAATCTGAAGCAGTTAAGCAATCAGAT 2178  
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp 200  
Qy 2179 CTTCTTCAGGTTATGACCATCTGTGCCAGTTTCGTAAATGTCTGTCACTTTCGACTCT 2238  
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220  
Qy 2239 GAGAACTCTCGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGCACAGACGA 2298  
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyValGlnGluTrpThrGluArg 240  
Qy 2299 CACGATATATAGTGGATGTCTAAAACGCATACATTT 2337  
Db 241 HisGlyTyrIleValAsnValSerLysArgIleProPhe 253

RESULT 7  
US-09-697-186B-15  
; Sequence 15, Application US/09697186B  
; Patent No. 6723543  
; GENERAL INFORMATION:  
; APPLICANT: YOKOYAMA, SIGEYUKI  
; APPLICANT: HOSEKI, JUN  
; APPLICANT: YANO, TAKATO  
; APPLICANT: KOYAMA, YOSHINORI  
; APPLICANT: KURAMITSU, SEIKI  
; APPLICANT: KAGAMIYAMA, HIROYUKI  
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD  
; TITLE OF INVENTION: OF SCREENING THERMOPHILIC BACTERIA USING THE SAME  
; FILE REFERENCE: 04853.0048-00000  
; CURRENT APPLICATION NUMBER: US/09/697,186B  
; CURRENT FILING DATE: 2000-10-27  
; PRIOR FILING DATE: 1999-10-29  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme  
; OTHER INFORMATION: obtained by introduction of point mutation into  
; OTHER INFORMATION: wild type KNT gene of Staphylococcus aureus and  
; OTHER INFORMATION: its expression  
US-09-697-186B-15

Alignment Scores:  
Pred. No.: 9.95e-131 Length: 253  
Score: 1268.00 Matches: 239  
Percent Similarity: 97.2% Conservative: 7  
Best Local Similarity: 94.5% Mismatches: 7  
Query Match: 23.0% Indels: 0  
DB: 2 Gaps: 0

US-10-681-086-1 (1-3156) x US-09-697-186B-15 (1-253)

Qy 1579 GTGAATGGACCAATAATACTAGACAGAGAAAGATGAGATTGTCATGAAATTAAG 1638  
Db 1938

```
Db      1 MetSerGlyProIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
Qy      1639 GAACCAATATGGATAATATGGGATGATGTTAAGGCTATTGGGTATTGGCTCTCTT 1698
Db      21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40
Qy      1699 GGTCTGTACAGCTGATGGCCCTATTCCGATATTGAGATGATGTGTCTCATGTCAACAGAG 1758
Db      41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetCysValMetSerThrGlu 60
Qy      1759 GAAGCAGAGTTCAGCCATGAATGGAACACCGGTGAGTGAAGGTGGAAGTGAATTTGAT 1818
Db      61 GlyAlaGluPheSerTyrGluTrpThrThrGlyGluTrpLysAlaGluValAsnPheTyr 80
Qy      1819 AGCGAAGAGATTCTACTAGATTATGCATCTCAGGTGGAATCAGATGCGCCCTTACACAT 1878
Db      81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluSerAspTrpProLeuThrHis 100
Qy      1879 GGTCAATTTTCTCTATTTCGCGATTATGATTCAGGTGGATATCTTAGAGAAAGTGTAT 1938
Db      101 GlyArgPhePheSerIleLeuProIleTyrAspProGlyGlyTyrPheGluLysValTyr 120
Qy      1939 CAACCTGTCTAAATCGGTAGACCCAAACGTTCCAGATGCGATTGTGCGCCCTTATCGTA 1998
Db      121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140
Qy      1999 GAAGAGCTGTTTGAATATGACGCAAAATGGCGTATATTCGTGTGTCAGAGGACCGACACA 2058
Db      141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
Qy      2059 TTCTTACCATTCTTCACTGTACAGGTAGCAATGGCAGGTGCGCATGTCATTGGTCTGCAT 2118
Db      161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
Qy      2119 CATCCATCTCTGTATACGACGAGCGCTTCGGTCTTAACTGAAGCAGTGTAAACAATCAGAT 2178
Db      181 HisArgIleCysTyrThrThrSerAlaThrValLeuThrGluAlaValLysLeuSerAsp 200
Qy      2179 CTTCCTTCAGGTATGACCACTGCGCAGTTCGTAATGTCGTGCACTTCCGACTCT 2238
Db      201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
Qy      2239 GAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTTCAGAGTGGACAGAACGA 2298
Db      221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyValGlnGlyTrpThrGluArg 240
Qy      2299 CACGGATATATAGTGGATGTGTCAAAACGCCATACCATTT 2337
Db      241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253
```

## RESULT 8

```
US-09-697-186B-2
; Sequence 2, Application US/09697186B
; Patent No. 6723543
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, SIGEYUKI
; APPLICANT: HOSEKI, JUN
; APPLICANT: YANO, TAKATO
; APPLICANT: KOYAMA, YOSHINORI
; APPLICANT: KURAMITSU, SEIKI
; APPLICANT: KAGAMIYAMA, HIROYUKI
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
; FILE REFERENCE: 04853.0048-00000
; CURRENT APPLICATION NUMBER: US/09/697,186B
; PRIORITY FILING DATE: 2000-10-27
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 253
; TYPE: PRT
```

```
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme
; OTHER INFORMATION: obtained by introduction of point mutation into
; OTHER INFORMATION: wild type KNT gene of Staphylococcus aureus and
; OTHER INFORMATION: its expression
US-09-697-186B-2
Alignment Scores: 1.28e-130 Length: 253
Pred. No.: 1267.00 Matches: 240
Score: 1267.00 Conservative: 4
Percent Similarity: 96.4% Mismatches: 9
Best Local Similarity: 94.9% Indels: 0
Query Match: 23.0% Gaps: 0
DB: 2
US-10-681-086-1 (1-3156) x US-09-697-186B-2 (1-253)
Qy      1579 GTGAATGCAACCAATAATAATGACTAGAGAGAAAGAAATGAAGATTGTTTCATGAAATTAAG 1638
Db      1 MetLysGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
Qy      1639 GAACCAATATGGATAATATATGCGGATGATGTTAAGGCTATTGGGTATTGGCTCTCTT 1698
Db      21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40
Qy      1699 GGTCTGTACAGCTGATGGCCCTATTCCGATATTGCGATATTCAGATGATGTGTGTCATGTCACAGAG 1758
Db      41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetCysValMetSerThrGlu 60
Qy      1759 GAACGACAGTTCAGCCATGAATGGAACACCGGTGAGTGAAGGTGGAAGTCAATTTTGCAT 1818
Db      61 GlyAlaGluPheSerTyrGluTrpThrThrGlyGluTrpLysAlaGluValAsnPheTyr 80
Qy      1819 AGCGAAGAGATTCTACTAGATTATGCATCTCAGGTGGAATCAGATGCGCCCTTATCGTA 1878
Db      81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluSerAspTrpProLeuThrHis 100
Qy      1879 GGTCAATTTTCTCTATTTCGCGATTATGATTCAGGTGGATATCTTAGAGAAAGTGTAT 1938
Db      101 GlyArgPhePheSerIleLeuProIleTyrAspProGlyGlyTyrPheGluLysValTyr 120
Qy      1939 CAACCTGTCTAAATCGGTAGACCCAAACGTTCCAGATGCGATTGTGCGCCCTTATCGTA 1998
Db      121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140
Qy      1999 GAAGAGCTGTTTGAATATGACGCAAAATGGCGTATATTCGTGTGTCAGAGGACCGACACA 2058
Db      141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
Qy      2059 TTTCCTACCATCTTCACTGTACAGGTAGCAATGGCAGGTGCGCATGTCATTGGTCTGCAT 2118
Db      161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
Qy      2119 CATCCATCTCTGTATACGACGAGCGCTTCGGTCTTAACTGAAGCAGTGTAAACAATCAGAT 2178
Db      181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp 200
Qy      2179 CTTCCTTCAGGTATGACCACTCTCTGCGAGTTCGTAATGTCGTGCACTTCCGACTCT 2238
Db      201 LeuProSerGlyTyrAspHisLeuCysGlnLeuValMetSerGlyGlnLeuSerAspSer 220
Qy      2239 GAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTTCAGAGTGGACAGAACGA 2298
Db      221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
Qy      2299 CACGGATATATAGTGGATGTGTCAAAACGCCATACCATTT 2337
Db      241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253
RESULT 9
US-09-697-186B-13
; Sequence 13, Application US/09697186B
```



```
QY 1819 AGCGAAGAGATCTACTAGATTATGCATCTCAGGTGGGAATCAGATTGGCCGCTTACACAT 1878
DB 81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluProAspTrpProLeuThrHis 100
QY 1879 GGTCAATTTTCTCTATTATTTCCGATTTATGATTCAGGTGGGATATCTAGAGAAAGTGTAT 1938
DB 101 GlyIysPhePheSerIleLeuProIleTyrAspProGlyGlyTyrLeuGluIysValTyr 120
QY 1939 CAAACTGCTAAATCGGTAGAACCCAAACGTTCCACGATGGGATTTGGCCCTTATCGTA 1998
DB 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140
QY 1999 GAAGAGCTGTTGTAATATGACGAGCGCTTCGGTCTTAATTCGTGTCAGAGGACCGACACAA 2058
DB 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
QY 2059 TTTCTACCATCTTGCATCTGACGTAGCAATGCGAGGTGCATGTTGATTTGGTCTGCAT 2118
DB 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
QY 2119 CATCGCATCTGTTATACAGACGAGCGCTTCGGTCTTAATTCGTAAGCAAGTTAAGCAATCAGAT 2178
DB 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp 200
QY 2179 CTTCCTTCAGGTATGACCATCTGTCAGTTCGTAATGTCGTCAACTTTCCGACTCT 2238
DB 201 LeuProSerGlyTyrValGlnLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
QY 2239 GAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGA 2298
DB 221 GluIysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
QY 2299 CACGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337
DB 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 11
US-09-697-186B-17
; Sequence 17, Application US/09697186B
; Patent No. 6723543
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, SIGEYUKI
; APPLICANT: HOSEKI, JUN
; APPLICANT: YANO, TAKATO
; APPLICANT: KAYAMA, YOSHINORI
; APPLICANT: KURAMITSU, SEIKI
; APPLICANT: KAGAMIYAMA, HIROYUKI
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
; FILE REFERENCE: 04853.0048-00000
; CURRENT APPLICATION NUMBER: US/09/697,186B
; CURRENT FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: JP 309616/1999
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme
; OTHER INFORMATION: obtained by introduction of point mutation into
; OTHER INFORMATION: wild type KNT gene of Staphylococcus aureus and
; OTHER INFORMATION: its expression
US-09-697-186B-17

Alignment Scores:
Pred. No.: 4,58e-130 Length: 253
Score: 1262.00 Matches: 240
Percent Similarity: 96.4% Conservative: 4
Best Local Similarity: 94.9% Mismatches: 9
```

```
Query Match: 22.9% Indel: 0
DB: 2 Gaps: 0
US-10-681-086-1 (1-3156) x US-09-697-186B-17 (1-253)
QY 1579 GTGAATGGACCAATAATAATGACTAGAGAAAGAAATGAAGATTGTTCAATAAATTAAG 1638
DB 1 MetaAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
QY 1639 GAACGAATATTTGATAAATATGGGATGATGTTAAGGCTATTTGGTGTATTTGGCTCTCTT 1698
DB 21 GluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyrGlySerLeu 40
QY 1699 GGTGCTCAGACTGATGGCCCTATTTCCGATATTCAGATGATGTGTGTCATGTCAACACAG 1758
DB 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60
QY 1759 GAACGACAGATTCCACCATGAATGGACAAACCGGTGAGTGGAGGTGAAGTGAATTTTGTAT 1818
DB 61 GlyAlaGluPheSerTyrGluTrpThrThrGlyGluTrpLysAlaGluValAsnPheTyr 80
QY 1819 AGCGAAGAGATCTACTAGATTATGCATCTCAGGTGGGAATCAGATTGGCCGCTTACACAT 1878
DB 81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluSerAspTrpProLeuThrHis 100
QY 1879 GGTCAATTTTCTCTATTATTTCCGATTTATGATTCAGGTGGATCTTAGAGAAAGTGTAT 1938
DB 101 GlyArgPhePheSerIleLeuProIleTyrAspProGlyGlyTyrPheGluIysValTyr 120
QY 1939 CAAACTGCTAAATCGGTAGAACCCAAACGTTCCACGATGGGATTTGGCCCTTATCGTA 1998
DB 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140
QY 1999 GAAGAGCTGTTGTAATATGACGAGCGCTTCGGTCTTAATTCGTGTCAGAGGACCGACACAA 2058
DB 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
QY 2059 TTTCTACCATCTTGCATCTGACGTAGCAATGCGAGGTGCATGTTGATTTGGTCTGCAT 2118
DB 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
QY 2119 CATCGCATCTGTTATACAGACGAGCGCTTCGGTCTTAATTCGTAAGCAAGTTAAGCAATCAGAT 2178
DB 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp 200
QY 2179 CTTCCTTCAGGTATGACCATCTGTCAGTTCGTAATGTCGTCAACTTTCCGACTCT 2238
DB 201 LeuProSerGlyTyrValGlnLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
QY 2239 GAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGA 2298
DB 221 GluIysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
QY 2299 CACGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337
DB 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 12
US-09-697-186B-18
; Sequence 18, Application US/09697186B
; Patent No. 6723543
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, SIGEYUKI
; APPLICANT: HOSEKI, JUN
; APPLICANT: YANO, TAKATO
; APPLICANT: KAYAMA, YOSHINORI
; APPLICANT: KURAMITSU, SEIKI
; APPLICANT: KAGAMIYAMA, HIROYUKI
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
; FILE REFERENCE: 04853.0048-00000
; CURRENT APPLICATION NUMBER: US/09/697,186B
; CURRENT FILING DATE: 2000-10-27
```



```

; PRIOR APPLICATION NUMBER: JP 309616/1999
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 253
;

```

Alignment Scores:			
Pred. No.:	4.58e-130	Length:	253
Score:	1362.00	Matches:	240
Percent Similarity:	96.4%	Conservative:	4
Best Local Similarity:	94.9%	Mismatches:	9
Query Match:	22.9%	Indels:	0
DB:	2	Gaps:	0

US-10-681-086-1 (1-3156) x US-09-697-186B-18 (1-253)

1579	GTGAATGGACCAATAATAATGACTAGAGAGAAGAATAAGATGTTCATGAAATTAAG	1638
QY	::::	
Db	1 MetAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys	20
QY	1639 GAACGAATATTGGATAAATATGGGATGATGTTAAGGCTATTGGTGTGTTTATGGCTCTC	1698
Db	21 GluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyrGlySerLeu	40
QY	1699 GGTTCGTCCAGACTCATGGGCCCTATTTCGGATATTGAGATGATGTGTGCATGTCAACAGAG	1758
Db	41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu	60
QY	1759 GAACGACAGATTCACGCCATGAATGACAAACCGGTGAGTGGAAAGTGGAAATTTTGAT	1818
Db	61 GlyAlaGluPheSerTyrGluThrThrThrGlyGluTrpLysAlaGluValAsnPheTyr	80
QY	1819 AGCGAAGAGATTCTACTAGATTATGCACTCAGGTGGAATCAGATTGGCCGCTTACACAT	1878
Db	81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluSerAspTrpProLeuThrHis	100
QY	1879 GGTCAATTTTCTCTATTTTCCGATTATGATTTCAGTGGATCTTAGAGAAGTGTAT	1938
Db	101 GlyArgPhePheSerIleLeuProIleTyrAspProGlyGlyTyrPheGluLysValTyr	120
QY	1939 CAAACTGCTAAATCGGTAGAGGCCAAACCGTCCACGATCGAATTTGTGCGCTTATCGTA	1998
Db	121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal	140
QY	1999 GAAGAGCTGTTTGAATATGCAGGCAAAATGCGCTAATATTCGTGTGCAAGCACCAACA	2058
Db	141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr	160
QY	2059 TTTCTACCATCTTGACTGTACAGTAGCAATGGCAGGTGCCATGTTGATGTGCTGCGAT	2118
Db	161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis	180
QY	2119 CATTCGCATCTGTTATACGACGAGCGCTTCGGTCTTTAACTCAAGCAGGTTAAGCAATCAGAT	2178
Db	181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp	200
QY	2179 CTTTCCTTCAGGTTATGACCATCTGTGCCAGTTTCGTAAATGTCTGGTCAACTTTCCGACTCT	2238
Db	201 LeuProSerGlyTyrValGlnLeuLysGlnPheValMetSerGlyGlnLeuSerAspSer	220
QY	2239 GAGAAATCTTCGGAATCGCTAGAGAAATTTCTGGAAATGGGAATTCGAGTAGTGGACAGACGA	2298
Db	221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg	240

Qy	2299	CACGATATATAGTGGATGTGTCAAAACGCATACCATT	2337
Db	241	HISGLYTYRILEVALASPVALSERLYSARGILEPROPHE	253

RESULT 13

US-09-697-186B-16  
; Sequence 16, Application US/09697186B  
; Patent No. 6723543

```
?
? APPLICANT: KAGAWAYAMA, HIROKUNI
? TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
? FOR SCREENING THERMOPHILIC BACTERIA USING THE SAME
? FILE REFERENCE: 04853-0048-00000
? CURRENT APPLICATION NUMBER: US/09/697,186B
? CURRENT FILING DATE: 2000-10-27
? PRIOR APPLICATION NUMBER: JP 309616/1999
? PRIOR FILING DATE: 1999-10-29
? NUMBER OF SEQ ID NOS: 20
? SOFTWARE: Patent In Ver. 2.1
?
```

Alignment Scores:					
Pred. No.:	7.62e-130	Length:	253		
Score:	1360.00	Matches:	237		
Percent Similarity:	96.0%	Conservative:	6		
Best Local Similarity:	93.7%	Mismatches:	10		
Query Match:	22.8%	Indels:	0		
DB:	2	Gaps:	0		

US-10-681-086-1 (1-3156) X US-09-697-186B-16 (1-253)

1579	Qy	GTCAATGGACCAATAATAATGACTAGAGAGAAGAAGATGAAGATTGTTTCATGAATAATTAAG	1638
	Db	::: ::::	20
1	MetAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys		
1639	Qy	GAACGAATATTGCATAAATAATGGGGATGATGTTAAAGCTATTGGTGTGTTTATGGCTCTCTT	1698
	Db	::: ::::	40
21	GluArgGileuAspLysTyrGlyArgAspValLysAlaIleGlyValTyrGlySerLeu		
1699	Qy	GGTCGTGCAGACTCATGGGCCCTATTCCGATATTAGAGATGATGTGTGCATGTCAACACAGAG	1758
	Db	::: ::::	60
41	GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu		
1759	Qy	GAAGCAGAGTTCAGCCATGAATCGACAACCGGTGAGTGGAGGTGGAGTGAATTTTGTAT	1818
	Db	::: ::::	80
61	GluThrGluPheSerTyrGluThrThrGlyGluTyrLysAlaGluValAsnPheTyr		
1819	Qy	ACCGAGAGATTCTACTAGATTATGCATCTCAGGTGGAATCAGATTGGCCGCTTACACAT	1878
	Db	::: ::::	100
81	SerGluGluIleuIleuAspTyrAlaSerArgValGluProAspTrpProLeuThrHis		
1879	Qy	GGTCAATTTTTTCTATTTCGCGATTATGATTACGGTGAGTACTTAGAGAAGTGATAT	1938
	Db	::: ::::	120
101	GlyLysPhePheSerIleuProIleTyrAspThrGlyGlyTyrLeuGluLysValTyr		
1939	Qy	CAAACTGCTAAATCGGTAGAGGCCCAAAAGTTCACAGATCGGATTTGTGCGCTTATCGTA	1998
	Db	::: ::::	140
121	GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal		

Qy 1999 GAAGAGCTGTTTGAATATGCAAGCAATCGGCTTAATATTCGTGCAAGGACCGACAACA 2058  
Db 141 GluGluLeuPheGluTyAlaGlyLysTrpArgAniLeArGValGlnGlyProLeuThr 160  
Qy 2059 TTTCTACCATCTTGACTGTACAGGTAGCAATGCGAGTGCATGTTGATGGTCTGCAT 2118  
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180  
Qy 2119 CATCGCATCTTTATACCAAGCGCTTCGGCTCTTAACGTGAAGCAATCAGAT 2178  
Db 181 HisArgIleCysTyThrThrGlyAlaSerValLeuThrGluAlaValArgGlnProAsp 200  
Qy 2179 CTTCCTTCAGGTTATGACCATCTGTGCGAGTTTCGTAATGCTCGGTCAACTTTCGACACT 2238  
Db 201 LeuProGlyTyAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220  
Qy 2239 GAGAACTTCCTGGAATCCTGAGAAATTCCTGGAATGCGATTCAGGAGTGACAGAAACA 2298  
Db 221 GluLysLeuLeuGluSerLeuGluAenPheTrpAenGlyIleGlnGluTrpAlaGluArg 240  
Qy 2299 CACGGATATAGTGATGTGTCAAAACGCATACCATTT 2337  
Db 241 HisGlyTyIleValAspValSerLysArgIleProPhe 253

## RESULT 14

US-09-697-186B-19  
; Sequence 19, Application US/09697186B  
; Patent No. 6723543  
; GENERAL INFORMATION:  
; APPLICANT: YOKOYAMA, SIGEYUKI  
; APPLICANT: HOSEKI, JUN  
; APPLICANT: YANO, TAKATO  
; APPLICANT: KOYAMA, YOSHINORI  
; APPLICANT: KURAMITSU, SEIKI  
; APPLICANT: KAGAMIYAMA, HIROYUKI  
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD  
; FILE REFERENCE: 04853.0048-00000  
; CURRENT APPLICATION NUMBER: US/09/697,186B  
; CURRENT FILING DATE: 2000-10-27  
; PRIOR APPLICATION NUMBER: JP 309616/1999  
; PRIOR FILING DATE: 1999-10-29  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 19  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme  
; OTHER INFORMATION: Obtained by introduction of point mutation into  
; OTHER INFORMATION: wild type Knt gene of Staphylococcus aureus and  
; OTHER INFORMATION: its expression  
US-09-697-186B-19

Alignment Scores:  
Pred. No.: 2,11e-129 Length: 253  
Score: 1256.00 Matches: 239  
Percent Similarity: 96.0% Conservative: 4  
Best Local Similarity: 94.5% Mismatches: 10  
Query Match: 22.8% Indels: 0  
DB: 2 Gaps: 0

US-10-681-086-1 (1-3156) x US-09-697-186B-19 (1-253)

Qy 1579 GTGAATGGACCAATAATGACTAGAGAAGAAATGAAGATGTTTCATGAATAAAG 1638  
Db 1 MetaAnGlyProIleLeuMetThrArgGluGluArgMetLysIleValHisGluIleLys 20  
Qy 1639 GAACGAATATGGATAAATATGGGATGATGTTAAGGCTATGTTGTTTATGGCTCTCTT 1698  
Db 21 GluArgIleLeuAspLysTyGlyAspValLysAlaIleGlyValTyGlySerLeu 40

Qy 1699 GGTCTGCAGACTGATGGCCCTTATTCGATATTGAGATGATGTGTCTCATGTCAACAGAG 1758  
Db 41 GlyArgGlnThrAspGlyProTySerAspIleGluMetMetCysValMetSerThrGlu 60  
Qy 1759 GAACAGAGTTTCACCCATGAATGACCAACCGGTGAGTGGAAAGTGGAAAGTGAATTTGAT 1818  
Db 61 GlyAlaGluPheSerTyGluTrpThrThrGlyGluTrpLysAlaGluValAsnPheTy 80  
Qy 1819 AGCCAAGAGATCTTACTAGATTATGATCTCAGGTGGATCAGATTCGCGCCCTTACACAT 1878  
Db 81 SerGluGluIleLeuLeuAspTyAlaSerArgValGluProAspTrpProLeuThrHis 100  
Qy 1879 GGTCAATTTTCTCTATTTTGGCCGATTTATGATTCAGGTGGATATCTTAGAAGAAAGTGTAT 1938  
Db 101 GlyLysPhePheSerIleLeuProIleTyAspProGlyGlyTyLeuGluLysValTy 120  
Qy 1939 CAAACTCTAAATCGGTAGAACCCAAACGTTCCACGATGCGATTTGTCCTTATTCGTA 1998  
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140  
Qy 1999 GAGAGCTGTTTGAATATGCAAGCAATGCGTAATATTCGTGTGCAAGGACCGACAACA 2058  
Db 141 GluGluLeuPheGluTyAlaGlyLysTrpArgAniLeArGValGlnGlyProThrThr 160  
Qy 2059 TTTCTACCATCTTGACTGTACAGGTAGCAATGCGAGTGCATGTTGATTCGTCTGCAT 2118  
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180  
Qy 2119 CATCGCATCTGTTATACGAGCGCTTCGGTCTTAACGTGAAGCAATCAGATCAGAT 2178  
Db 181 HisArgIleCysTyThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp 200  
Qy 2179 CTTCCTTCAGGTTATGACCATCTGTGCGAGTTTCGTAATGCTGTGCTCAACTTTCGACTCT 2238  
Db 201 LeuProSerGlyTyValGlnLeuCysGlnPheValMetSerGlyGlnLeuSerAspPro 220  
Qy 2239 GAGAACTTCCTGGAATCGCTAGAGAATTCCTGGAATGCGATTCAGGAGTGACAGAAACA 2298  
Db 221 GluLysLeuLeuGluSerLeuGluAenPheTrpAenGlyIleGlnGluTrpThrGluArg 240  
Qy 2299 CACGGATATATAGTGATGTGTCAAAACGCATACCATTT 2337  
Db 241 HisGlyTyIleValAspValSerLysArgIleProPhe 253

## RESULT 15

US-09-697-186B-3  
; Sequence 3, Application US/09697186B  
; Patent No. 6723543  
; GENERAL INFORMATION:  
; APPLICANT: YOKOYAMA, SIGEYUKI  
; APPLICANT: HOSEKI, JUN  
; APPLICANT: YANO, TAKATO  
; APPLICANT: KOYAMA, YOSHINORI  
; APPLICANT: KURAMITSU, SEIKI  
; APPLICANT: KAGAMIYAMA, HIROYUKI  
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD  
; FILE REFERENCE: 04853.0048-00000  
; CURRENT APPLICATION NUMBER: US/09/697,186B  
; CURRENT FILING DATE: 2000-10-27  
; PRIOR APPLICATION NUMBER: JP 309616/1999  
; PRIOR FILING DATE: 1999-10-29  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme  
; OTHER INFORMATION: Obtained by introduction of point mutation into  
; OTHER INFORMATION: wild type Knt gene of Staphylococcus aureus and  
; OTHER INFORMATION: its expression

US-09-697-186B-3

Alignment Scores:

Pred. No.:	1.21e-125	Length:	253
Score:	1222.00	Matches:	231
Percent Similarity:	93.7%	Conservative:	6
Best Local Similarity:	91.3%	Mismatches:	16
Query Match:	22.1%	Indels:	0
DB:	2	Gaps:	0

US-10-681-086-1 (1-3156) x US-09-697-186B-3 (1-253)

Qy	1579	GTGAATCGACCAATATAATGACTAGAGAGAAAGATGATGTTCAATGAATTAAG	1638
Db	:	:	:
Db	1	MetLysGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys	20
Qy	1639	GAACGAATATGGATAATATGGGATGATGTTAAGGCTATTGGTGTATTGGCTCTCTT	1698
Db	21	GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu	40
Qy	1699	GGTCGTCAGACTGATGGCCCTATTTCGATATTGAGATGATGTGTGTCATGTCAACAGAG	1758
Db	41	GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValLeuSerThrGlu	60
Qy	1759	GAAGCAGAGTTCAGCCATGAATGACCAACCGGTGAGTGGAGGTGGAAGTCAATTTTGAT	1818
Db	61	GlyValGluPheSerTyrGluTyrThrThrGlyGluTyrLysAlaGluValAsnPheTyr	80
Qy	1819	AGCGAAGAGATTCCTAGATTTATGCTATCTCAGGTGAATCAGATTCGGCGCTTACACAT	1878
Db	81	SerGluGluIleLeuLeuAspTyrAlaSerArgValGluProAspTyrProLeuThrHis	100
Qy	1879	GGTCAATTTTCTCTATTTCGCGATTTATGATTTCAGTGGATCTTAGAGAAAGTGTAT	1938
Db	101	GlyArgPhePheSerIleLeuProIleTyrAspProGlyGlyTyrPheGluLysValTyr	120
Qy	1939	CAAACTGCTAAATCGGTAGAGCCCAACGTTCCACGATGCGATTTGTGCCCTTTATCGTA	1998
Db	121	GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal	140
Qy	1999	GAAGAGCTGTTTGAATATGCGGCAATATGGCGTAATATTGTTGTGCAAGGACGACAACA	2058
Db	141	GluGluLeuPheGluTyrAlaGlyLysTyrArgAsnIleArgValGlnGlyProThrThr	160
Qy	2059	TTTCTACCATCTTGACTGTACAGTAGCAGTGCATGCCATGTCATGTTGTTGTCGTCAT	2118
Db	161	PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis	180
Qy	2119	CATCGCATCTGTTATACGACGAGCGCTTCGGTCTTTAACTGAAGCAGTTAAGCAATCAGAT	2178
Db	181	HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp	200
Qy	2179	CTTCCTTCAGTTATGACCATCTGTGCGAGTTCGTAATGCTGTGTCAACTTCCGACTCT	2238
Db	201	LeuProGlyTyrValGlnLeuCysGlnLeuValMetSerGlyGlnLeuSerAspPro	220
Qy	2239	GAGAACTTCTGGAATCGTAGAATTTCTGGAATTCGGAATTCAGGAGTGGACAGACGA	2298
Db	221	GluLysLeuLeuGluSerLeuGluAsnPheTyrAsnGlyValGlnGluTyrPalGluArg	240
Qy	2299	CACGGATATATAGTGGATGTGTCAAAACGCATACCATTT	2337
Db	241	HisGlyTyrIleValAspValSerLysArgIleProPhe	253

Search completed: February 10, 2006, 03:03:33  
Job time : 111 secs

***This Page Blank (uspto)***

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: February 10, 2006, 03:49:13 ; Search time 87.9 Seconds  
(without alignments)  
3000.385 Million cell updates/sec

Title: us-10-681-086-1

Perfect score: 5517

Sequence: 1 ggaatccagggattacagc.....atcggtatgatgcgaattc 3156

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 3735138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp  
-O=/abs/ABSSWEB\_spool/US10681086/runat\_09022006\_140527\_11976/app\_query.fasta\_1  
-DB=Published Applications\_AA\_Main -QWMT=fastan -SUFFIX=n2p.rapbm  
-MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1  
-MATRIX=blosum62 -TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct  
-THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext  
-HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abes02p  
-USER=US10681086 @CGN 1.1 307 @runat\_09022006\_140527\_11976 -NCPU=6 -ICPU=3  
-NO\_WMAP -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA Main:

1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubppaa/US08\_PUBCOMB.pep:\*  
3: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
4: /cgn2\_6/ptodata/1/pubppaa/US10A\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubppaa/US10B\_PUBCOMB.pep:\*  
6: /cgn2\_6/ptodata/1/pubppaa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1836	33.3	389	4	US-10-369-493-23255
2	1333	24.2	253	4	US-10-320-800-75
3	1333	24.2	253	4	US-10-718-628-11
4	1318	23.9	253	4	US-10-718-628-1
5	1282	23.1	253	4	US-10-718-628-12
6	1277	23.2	253	4	US-10-718-628-14
7	1268	23.0	253	4	US-10-718-628-15
8	1267	23.0	253	4	US-10-718-628-2
9	1267	23.0	253	4	US-10-718-628-13
10	1265	22.9	253	4	US-10-718-628-20
11	1262	22.9	253	4	US-10-718-628-17

12	1262	22.9	253	4	US-10-718-628-18
13	1260	22.8	253	4	US-10-718-628-16
14	1256	22.8	253	4	US-10-718-628-19
15	1222	22.1	253	4	US-10-718-628-3
16	942	17.1	387	4	US-10-033-078-6
17	942	17.1	387	4	US-10-763-333-6
18	851	15.4	338	4	US-10-033-078-8
19	851	15.4	338	4	US-10-763-933-8
20	828	15.3	332	4	US-10-369-493-23135
21	828	15.0	321	4	US-10-282-122A-71209
22	828	15.0	327	4	US-10-724-972A-4421
23	809	14.7	332	4	US-10-282-122A-45833
24	793	14.4	338	3	US-09-815-242-5784
25	793	14.4	335	3	US-09-815-242-12780
26	793	14.4	335	3	US-09-815-242-13125
27	793	14.4	335	4	US-10-282-122A-44022
28	778	14.1	321	4	US-10-282-122A-71676
29	767	13.9	371	4	US-10-369-493-22995
30	712	12.9	392	4	US-10-369-493-13494
31	712	12.9	395	4	US-10-369-493-17471
32	711	12.9	396	4	US-10-282-122A-46084
33	710	12.9	388	4	US-10-369-493-16649
34	705	12.8	395	3	US-09-815-242-12441
35	704	12.8	385	3	US-09-815-242-5575
36	704	12.8	396	4	US-10-282-122A-58006
37	700	12.7	393	4	US-10-369-493-18251
38	699	12.7	324	4	US-10-282-122A-45547
39	699	12.7	398	4	US-10-282-122A-71674
40	699	12.7	400	4	US-10-369-493-20208
41	698.5	12.7	372	4	US-10-369-493-1058
42	695.5	12.6	400	5	US-10-501-282-2370
43	692	12.5	398	4	US-10-369-493-21684
44	689	12.5	391	4	US-10-369-493-16710
45	688	12.5	402	4	US-10-369-493-18774

#### ALIGNMENTS

##### RESULT 1

US-10-369-493-23255  
; Sequence 23255, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 23255  
; LENGTH: 389  
; TYPE: PRT  
; ORGANISM: Bacillus subtilis  
; US-10-369-493-23255

Alignment Scores:  
Pred. No.: 3 6e-173 Length: 389  
Score: 1836.00 Matches: 359  
Percent Similarity: 100.0% Conservative: 1  
Best Local Similarity: 99.7% Mismatches: 0  
Query Match: 33.3% Indels: 0  
DB: 4 Gaps: 0

US-10-681-086-1 (1-3156) x US-10-369-493-23255 (1-389)

QY 22 TTGAAGATTGATTCCTGGTTAAACGAGCGGTAGACAGATGAAGAAGCCGCGTACAT 81

```
Db      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
1  MetLysIleAspSerTrpLeuAsnGluArgLeuAspArgMetLysGluAlaGlyValHis 20
QY      82  CGTAACCTGCGGTCAATCGATGGAGCCGGTTCAGAGAGAAATATTGATGGCGAAAT 141
Db      21  ArgAsnLeuArgSerMetAspGlyAlaProValProGluArgAsnIleAspGlyGluAsn 40
QY      142  CAAACGGTCTGGTCTCTCAAACAATATTATTAGGGCTCGCAAGCGATAGACGTTTGATCGAT 201
Db      41  GlnThrValTrpSerSerAsnAsnTyrLeuGlyLeuAlaSerAspArgArgLeuIleAsp 60
QY      202  GCAGCCCAACAGCATTCAGCAATTTTCGGGACAGGAACGACGGTTCACGTTTAAACGACA 261
Db      61  AlaAlaGlnThrAlaLeuGlnGlnPheGlyThrGlySerSerGlySerArgLeuThrThr 80
QY      262  GGCAATTCGGTCTGGCATGAAAGCTAGAAAGATTCAGAAAGATTCAGCTTTAAACTGCAGAA 321
Db      81  GlyAsnSerValTrpHisGluLysLeuGluLysIleAlaSerPheLysLeuThrGlu 100
QY      322  CGGGCCCTGCTGTTTTCGAGCGGTTACTTGGCCAAATGTCGGTGTCTTTCATCTCTTGCCA 381
Db      101  AlaAlaLeuLeuPheSerSerGlyTyrLeuAlaAsnValGlyValLeuSerSerLeuPro 120
QY      382  GAAAGGAGATGTCATTTTAAGTGACCAAGCTCAATCATGCAAGTATGATCGACGGCTGC 441
Db      121  GluLysGluAspValIleLeuSerAspGlnLeuAsnHisAlaSerMetIleAspGlyCys 140
QY      442  CGACTTTCAGGCTGATACAGTGTGTTTATCGGCATATTGATATGATGATGATCTTGAAC 501
Db      141  ArgLeuSerLysAlaAspThrValValTyrArgHisIleAspMetAsnAspLeuGluAsn 160
QY      502  AAGCTGAATGAAACACACAGCGTTATCAGCGCGTTCATCGTAAACAGCGAGTATTCAGC 561
Db      161  LysLeuAsnGluThrGlnArgTyrGlnArgArgPheIleValThrAspGlyValPheSer 180
QY      562  ATGATGACCAATCGCCCTCTTGATCAGATCATCTCACTTGGCAACGCTATCATGCC 621
Db      181  MetAspGlyThrIleAlaProLeuAspGlnIleSerLeuAlaLysArgTyrHisAla 200
QY      622  TTCGTGTCGTGATGATGCCCAACAGAGAGTTTGGCGATTCGGGACAGGAACG 681
Db      201  PheValValValAspAspAlaHisAlaThrGlyValLeuGlyAspSerGlyGlnGlyThr 220
QY      682  AGTGAATACTTTGGTGTTCGCCACATTTGTAATCGGCACCTTAAGCAAAAGCTGTGGC 741
Db      221  SerGluTyrPheGlyValCysProAspIleValIleGlyThrLeuSerLysAlaValGly 240
QY      742  CGGAGGAGGTTTTCGGCAGGATTCACGGTCTTCATCGACTTTTCTGCTGAACCATGCC 801
Db      241  AlaGluGlyGlyPheAlaAlaGlySerAlaValPheIleAspPheLeuLeuAsnHisAla 260
QY      802  AGAACATTTATCTTCAACCGCTATTCGGCCAGCCAGCTGTGGCTGCTCACGAGCT 861
Db      261  ArgThrPheIlePheGlnThrAlaIleProProAlaSerCysAlaAlaAlaHisGluAla 280
QY      862  TTCACATCATTTGAAGCCAGGAGGAAACGACAGCTTTTATTTCTATATCAGCATG 921
Db      281  PheAsnIleIleGluAlaSerArgGluLysArgGlnLeuLeuPheSerTyrIleSerMet 300
QY      922  ATCAGAACCAAGTCTGAAGAATATGGTTATGTGGTGAAGGAGATCACACCCGATTTAT 981
Db      301  IleArgThrSerLeuLysAsnMetGlyTyrValValLysGlyAspHisThrProIleIle 320
QY      982  CCTGTAGTCATTTGGCGATGCCATAAACGGTCTCTATTTGCTGAAAAACTCGAGGCAAG 1041
Db      321  ProValValIleGlyAspAlaHisLysThrValLeuPheAlaGluLysLeuGlnGlyLys 340
QY      1042  GGAATTTATGCTCCCTGCAATTCGGCCGCAACCGTTGCGCGGGTGAAGCCGATTCGA 1101
Db      341  GlyIleTyrAlaProAlaIleArgProProThrValAlaProGlyGluSerArgIleArg 360
```

RESULT 2

US-10-320-800-75

```
Sequence 75, Application US/10320800
Publication No. US20030215469A1
GENERAL INFORMATION:
APPLICANT: ROBINSON, ANDREW
APPLICANT: GORRINGE, ANDREW
APPLICANT: HUDSON, MICHAEL
APPLICANT: REDDIN, KAREN
TITLE OF INVENTION: MULTICOMPONENT MENINGOCOCCAL VACCINE
FILE REFERENCE: 1581.0790001
CURRENT APPLICATION NUMBER: US/10/320,800
CURRENT FILING DATE: 2002-12-17
PRIOR APPLICATION NUMBER: PCT/GB99/03626
PRIOR FILING DATE: 1999-11-02
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn version 3.1
SEQ ID NO 75
LENGTH: 253
TYPE: PRT
ORGANISM: unidentified
FEATURE:
NAME/KEY: MISC FEATURE
OTHER INFORMATION: kanamycin resistance
US-10-320-800-75

Alignment Scores:
Pred. No.: 3,78e-123 Length: 253
Score: 1333.00 Matches: 252
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.6% Mismatches: 0
Query Match: 24.2% Indels: 0
DB: Gaps: 0

US-10-681-086-1 (1-3156) x US-10-320-800-75 (1-253)
QY 1579 GTGAATCGCAATAATAATGACTAGAGAGAAAGATGAAGATTGTTTCATGAATTAAG 1638
Db 1 MetAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
QY 1639 GAACGAATATTGGATAAATATGGGATGATGTTAAAGCTATTGCTGTTTATGGCTCTCTT 1698
Db 21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40
QY 1699 GGTGTCAGACTGATGGCCCTATTTCGGATATTTCAGATGATGTCATGTCACACAG 1758
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60
QY 1759 GAACGAGATTTCACCCATGAATGGACAAACCGTGAGTGGAAAGTGGAAAGTGAATTTGAT 1818
Db 61 GluAlaGluPheSerHisGluTyrThrThrGlyLysValGluValAsnPheAsp 80
QY 1819 AGCGAAGAGATTCTACTAGATTATGCAATCTCAGGTGGAATCAGATTGGCCCTTACACAT 1878
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpProLeuThrHis 100
QY 1879 GGTCAATTTTCTCTATTTTGGCCGATTATGATTTCAGTGGATACCTTAGAGAAAGTGAT 1938
Db 101 GlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120
QY 1939 CAAACTCCTAAATCGGTAGAGCCCAACGTTCCACAGTGGGATTTGGCCCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAlaLeuIleVal 140
QY 1999 GAACGAGCTGTTTGAATATGACGCAAAATGGCGTAATATTTCGTGTGCAAGGACCCAGCAACA 2058
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
QY 2059 TTTCTACCATCTCTTGACTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTGGTCTGCAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
QY 2119 CATCGATCTGTTATACGACGAGCGCTTCGGTCTTAAGTGAAGGAGTTAAGCAATCAGAT 2178
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200
```



```
Qy 1579 GTGAATGACCAATAATATGACTAGAGAGAAAGATGTTTCATCAAAATTAAG 1638
Db : : : : :
1 MetAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20

Qy 1639 GAACGAATATTGGATAAATATGAGGATGATGTTAAGGCTATTGGTGTATTGGCTCTCTT 1698
Db : : : : :
21 GluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyrGlySerLeu 40

Qy 1699 GGTCTCAGACTGATGGCCCTATTTCGATATTCAGATGATGTCATCTCAACAGAG 1758
Db : : : : :
41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetCysValMetSerThrGlu 60

Qy 1759 GAAGCAGAGTTCAGCCATTAATGACAAACCGTTCAGTGAAGGAGTCAATTTTGCAT 1818
Db : : : : :
61 GluAlaGluPheSerHisGluTrpThrGlyGluTrpLysValGluValAsnPheTyr 80

Qy 1819 AGCGAAGAGATTCCTACTAGATATGATCTCAGGTGGAATCAGATGGCCCTTACACAT 1878
Db : : : : :
81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpProLeuThrHis 100

Qy 1879 GGTCAATTTTCTCTATTTCGCTATTCAGTTCAGTGGATCTTACAGAAAGTGTAT 1938
Db : : : : :
101 GlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120

Qy 1939 CAAACTGCTAAATCGGTAGAACCCCAACGTTCCACGATGCGATTGTGCGCTTATCGTA 1998
Db : : : : :
121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140

Qy 1999 GAAGAGCTGTTGTAATATGACGAGCAAAATGGCGTAAATATTCGTGTGCAAGGCCGACAACA 2058
Db : : : : :
141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160

Qy 2059 TTTCTACCATCTTGACACTGACNGTAGCATGCGAGTGGCATGTTGATTTGCTGTCAT 2118
Db : : : : :
161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180

Qy 2119 CATCGACTCTGTATACGACGAGCGCTTCGGTCTTAACGAGCAGATTAAAGCAATTCAGAT 2178
Db : : : : :
181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200

Qy 2179 CTTCTCTCAGGTTATGACCAATTCGATGCGATTCGATGCGATTCGATGCGATTCGAT 2238
Db : : : : :
201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220

Qy 2239 GAGAACTCTTGGAAATCGCTAGAGATTCGATGCGATTCGATGCGATTCGATGCGATTCGAT 2298
Db : : : : :
221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240

Qy 2299 CACGGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337
Db : : : : :
241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253
```

RESULT 5

```
US-10-718-628-12
; Sequence 12, Application US/10718628
; Publication No. US2004007016A1
; GENERAL INFORMATION:
; APPLICANT: YOKOTAMA, SIGEYUKI
; APPLICANT: HOSEKI, JUN
; APPLICANT: YANO, TAKATO
; APPLICANT: KOYAMA, YOSHINORI
; APPLICANT: KURAMITSU, SEIKI
; APPLICANT: KAGAMIYAMA, HIROYUKI
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
; TITLE OF INVENTION: OF SCREENING THERMOPHILIC BACTERIA USING THE SAME
; FILE REFERENCE: 04853.0048-00000
; CURRENT APPLICATION NUMBER: US/10718,628
; CURRENT FILING DATE: 2003-11-24
; PRIOR APPLICATION NUMBER: US/09/697,186B
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: JP 309616/1999
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 20
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 12
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme
; OTHER INFORMATION: obtained by introduction of point mutation into
; OTHER INFORMATION: wild type KNT gene of Staphylococcus aureus and
; OTHER INFORMATION: its expression
US-10-718-628-12

Alignment Scores:
Pred. No.: 4.56e-118 Length: 253
Score: 1282.00 Matches: 241
Percent Similarity: 97.6% Conservative: 6
Best Local Similarity: 95.3% Mismatches: 6
Query Match: 23.2% Indels: 0
DB: Gaps: 0
```

US-10-681-086-1 (1-3156) x US-10-718-628-12 (1-253)

```
Qy 1579 GTGAATGACCAATAATATGACTAGAGAGAAAGATGTTTCATCAAAATTAAG 1638
Db : : : : :
1 MetAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValTyrGluIleLys 20

Qy 1639 GAACGAATATTGGATAAATATGAGGATGATGTTAAGGCTATTGGTGTATTGGCTCTCTT 1698
Db : : : : :
21 GluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyrGlySerLeu 40

Qy 1699 GGTCTCAGACTGATGGCCCTATTTCGATATTCAGATGATGTCATCTCAACAGAG 1758
Db : : : : :
41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetCysValMetSerThrGlu 60

Qy 1759 GAAGCAGAGTTCAGCCATTAATGACAAACCGTTCAGTGAAGGAGTCAATTTTGCAT 1818
Db : : : : :
61 GluAlaGluPheSerHisGluTrpThrGlyGluTrpLysAlaGluValAsnPheTyr 80

Qy 1819 AGCGAAGAGATTCCTACTAGATATGATCTCAGGTGGAATCAGATGGCCCTTACACAT 1878
Db : : : : :
81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluSerAspTrpProLeuThrHis 100

Qy 1879 GGTCAATTTTCTCTATTTCGCGATTTATGATTCAGTGGATCTTACAGAAAGTGTAT 1938
Db : : : : :
101 GlyArgPhePheSerIleLeuProIleTyrAspProGlyGlyTyrPheGluLysValTyr 120

Qy 1939 CAAACTGCTAAATCGGTAGAACCCCAACGTTCCACGATGCGATTGTGCGCTTATCGTA 1998
Db : : : : :
121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140

Qy 1999 GAAGAGCTGTTGTAATATGACGAGCAAAATGGCGTAAATATTCGTGTGCAAGGCCGACAACA 2058
Db : : : : :
141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160

Qy 2059 TTTCTACCATCTTGACACTGATGCGATTCGATGCGATTCGATGCGATTCGATGCGATTCGAT 2118
Db : : : : :
161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180

Qy 2119 CATCGATCTGTTATACGAGAGCGCTTCGCTCTTAACGAGCAGATTAAAGCAATTCAGAT 2178
Db : : : : :
181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaLeuLysGlnSerAsp 200

Qy 2179 CTTCTCTCAGGTTATGACCAATTCGATGCGATTCGATGCGATTCGATGCGATTCGATGCGATTCGAT 2238
Db : : : : :
201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220

Qy 2239 GAGAACTCTTGGAAATCGCTAGAGATTCGATGCGATTCGATGCGATTCGATGCGATTCGAT 2298
Db : : : : :
221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyValGlnGluTrpAlaGluArg 240

Qy 2299 CACGGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337
Db : : : : :
241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253
```



```
RESULT 6
US-10-718-628-14
; Sequence 14, Application US/10718628
; Publication No. US20040077016A1
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, SIGEYUKI
; APPLICANT: HOSEKI, JUN
; APPLICANT: YANO, TAKATO
; APPLICANT: KOYAMA, YOSHINORI
; APPLICANT: KURAMITSU, SEIKI
; APPLICANT: KAGAMIYAMA, HIROYUKI
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
; TITLE OF INVENTION: OF SCREENING THERMOPHILIC BACTERIA USING THE SAME
; FILE REFERENCE: 04853.0048-00000
; CURRENT APPLICATION NUMBER: US/10718,628
; CURRENT FILING DATE: 2003-11-24
; PRIOR FILING DATE: 2000-10-27
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme
; OTHER INFORMATION: obtained by introduction of point mutation into
; OTHER INFORMATION: wild type KNT gene of Staphylococcus aureus and
; OTHER INFORMATION: its expression
US-10-718-628-14
Alignment Scores:
Pred. No.: 1,44e-117 Length: 253
Score: 1277.00 Matches: 241
Percent Similarity: 97.6% Conservatives: 6
Best Local Similarity: 95.3% Mismatches: 6
Query Match: 23.1% Indels: 0
DB: 4 Gaps: 0
US-10-681-086-1 (1-3156) x US-10-718-628-14 (1-253)
QY 1579 GTGAATGGACCAATAAATGACTAGAGAGAAAGATGATGTTTCATGAATTAAG 1638
Db 1 MetAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
QY 1639 GAACGAATATTGGATAAATATGGGGATGATGTTAAGGCTATTGGTGTTCCTCTCT 1698
Db 21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40
QY 1699 GTGCTGACACTGATGGCCCTATTTCGGATATTGAGATGATGTTGTCATGTCACAGAG 1758
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60
QY 1759 GAACGAGAGTTTCAGCCATGAATGACACACCGTGAGTGAAGTGAAGTGAATTTGAT 1818
Db 61 GlyAlaGluPheSerTyrGluThrThrGlyGluTrpLysAlaGluValAsnPheTyr 80
QY 1819 AGCGAAGAGATTCTACTAGATTATGCACTCTCAGGTGGAATCAGATGGCCGCTTACACAT 1878
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluProAspTrpProLeuThrHis 100
QY 1879 GGTCAATTTTCTTATTTTCCGATTTATGATTTCAGTGGATCTTAGAGAAAGTGATAT 1938
Db 101 GlyArgPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120
QY 1939 CAACACTGTAATCGGTAGAGCCCAACGTTCCACCATGCGATTTCGTGCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140
QY 1999 GAAGAGCTGTTGTAATATGACGGCAAAATGGCGTAATATTCTGTGTGCAAGGACCAACA 2058
Db 21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
QY 2059 TTTCTACCATCTTGAGTACTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTGTTCTCAT 2118
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
QY 2119 CATCGCATCTGTTATACGACGAGCGCTTCGCTCTTAACCTGAAGCAGTGAACCAATCAGAT 2178
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp 200
QY 2179 CTTCTTCAGGTATGACCATCTGTGCCAGTTCGTAATGTCGTCTCACTTCGACTCT 2238
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
QY 2239 GAGAAACTCTCGAATCGCTAGAGAATTCTTGAATGGGATTCAGGAGTCGACAGACGA 2298
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyValGlnGluTrpThrGluArg 240
QY 2299 CACGGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337
Db 241 HisGlyTyrIleValAsnValSerLysArgIleProPhe 253
RESULT 7
US-10-718-628-15
; Sequence 15, Application US/10718628
; Publication No. US20040077016A1
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, SIGEYUKI
; APPLICANT: HOSEKI, JUN
; APPLICANT: YANO, TAKATO
; APPLICANT: KOYAMA, YOSHINORI
; APPLICANT: KURAMITSU, SEIKI
; APPLICANT: KAGAMIYAMA, HIROYUKI
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
; TITLE OF INVENTION: OF SCREENING THERMOPHILIC BACTERIA USING THE SAME
; FILE REFERENCE: 04853.0048-00000
; CURRENT APPLICATION NUMBER: US/10718,628
; CURRENT FILING DATE: 2003-11-24
; PRIOR FILING DATE: 2000-10-27
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme
; OTHER INFORMATION: obtained by introduction of point mutation into
; OTHER INFORMATION: wild type KNT gene of Staphylococcus aureus and
; OTHER INFORMATION: its expression
US-10-718-628-15
Alignment Scores:
Pred. No.: 1,13e-116 Length: 253
Score: 1268.00 Matches: 239
Percent Similarity: 97.2% Conservatives: 7
Best Local Similarity: 94.5% Mismatches: 7
Query Match: 23.0% Indels: 0
DB: 4 Gaps: 0
US-10-681-086-1 (1-3156) x US-10-718-628-15 (1-253)
QY 1579 GTGAATGGACCAATAAATGACTAGAGAGAAAGATGATGTTTCATGAATTAAG 1638
Db 1 MetSerGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
QY 1639 GAACGAATATTGGATAAATATGGGGATGATGTTAAGGCTATTGGTGTTCCTCTCT 1698
Db 21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40
QY 1699 GTGCTGACACTGATGGCCCTATTTCGGATATTGAGATGATGTTGTCATGTCACAGAG 1758
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60
QY 1759 GAACGAGAGTTTCAGCCATGAATGACACACCGTGAGTGAAGTGAAGTGAATTTGAT 1818
Db 61 GlyAlaGluPheSerTyrGluThrThrGlyGluTrpLysAlaGluValAsnPheTyr 80
QY 1819 AGCGAAGAGATTCTACTAGATTATGCACTCTCAGGTGGAATCAGATGGCCGCTTACACAT 1878
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluProAspTrpProLeuThrHis 100
QY 1879 GGTCAATTTTCTTATTTTCCGATTTATGATTTCAGTGGATCTTAGAGAAAGTGATAT 1938
Db 101 GlyArgPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120
QY 1939 CAACACTGTAATCGGTAGAGCCCAACGTTCCACCATGCGATTTCGTGCTTATCGTA 1998
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140
QY 1999 GAAGAGCTGTTGTAATATGACGGCAAAATGGCGTAATATTCTGTGTGCAAGGACCAACA 2058
Db 21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40
```

```
QY 1699 GGTCTGACAGTGGCCCTATTTCGGATATTGAGATGATGTGTGTCATGTCAACAGAG 1758
DB 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetCysValMetSerThrGlu 60

QY 1759 GAACGAGAGTTCAGCCATGAATGGACAACCGGTGAGTGAAGGTGGAAGTGAATTTGAT 1818
DB 61 GlyAlaGluPheSerTyrGluTrpThrThrGlyGluTrpLysAlaGluValAsnPheTyr 80

QY 1819 AGCGNAGAGATCTACTAGATTATGCATCTCAGTGGGAATCAGATTGGCCGCTTACAT 1878
DB 81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluSerAspTrpProLeuThrHis 100

QY 1879 GGTCAATTTTCTCTATTTCGCGATTATGATTCAGGTGGATATCTTAGAGAAAGTGTAT 1938
DB 101 GlyArgPhePheSerIleLeuProIleTyrAspProGlyGlyTyrPheGluLysValTyr 120

QY 1939 CAAACTGCTAAATCGGTAGAACCCCAACGGTTCACAGATGCGATTTGTCCTTATCGTA 1998
DB 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140

QY 1999 GAAGAGCTTTTGAATATGCAGGCAAAATGGCGTAATATTCGTGTCAGAGGACCGACA 2058
DB 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160

QY 2059 TTTCTACCATCTCTGACTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTGGTCTGCAT 2118
DB 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180

QY 2119 CATCGATCTGTTATPACAGCAGCGCTTCGGTCTTAATCAAGCAAGCAAGTAAAGCAAT 2178
DB 181 HisArgIleCysTyrThrThrSerAlaThrValLeuThrGluAlaValLysLeuSerAsp 200

QY 2179 CTTCCTTCAGGTATGACCATCTGCGCAGTTCGTAATGTCGTCACTTCGCACTCT 2238
DB 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220

QY 2239 GAGAACTCTCGGAATCGCTAGAGAAATTCGGAATGGGATTCAGGAGTGGACAGAACGA 2298
DB 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyValGlnGluTrpThrGluArg 240

QY 2299 CACGGATATATAGTGGATGTCTCAAAACGCATACCATTT 2337
DB 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253
```

## RESULT 8

```
US-10-718-628-2
; Sequence 2, Application US/10718628
; Publication No. US20040077016A1
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, SIGEYUKI
; APPLICANT: HOSEKI, JUN
; APPLICANT: YANO, TAKATO
; APPLICANT: KOYAMA, YOSHINORI
; APPLICANT: KURAMITSU, SEIKI
; APPLICANT: KAGAMIYAMA, HIROYUKI
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD
; OF SCREENING THERMOPHILIC BACTERIA USING THE SAME
; FILE REFERENCE: 04853.0048-00000
; CURRENT APPLICATION NUMBER: US/10/718,628
; CURRENT FILING DATE: 2003-11-24
; PRIOR APPLICATION NUMBER: US/09/697,186B
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: JP 309616/1999
; PRIOR FILING DATE: 1999-10-29
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 253
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme
```

```
; OTHER INFORMATION: obtained by introduction of point mutation into
; OTHER INFORMATION: wild type KOT gene of Staphylococcus aureus and
; OTHER INFORMATION: its expression
US-10-718-628-2
```

## Alignment Scores:

```
Pred. No.: 1.42e-116 Length: 253
Score: 1267.00 Matches: 240
Percent Similarity: 96.4% Conservative: 4
Best Local Similarity: 94.9% Mismatches: 9
Query Match: 23.0% Indels: 0
DB: 4 Gaps: 0
```

US-10-681-086-1 (1-3156) x US-10-718-628-2 (1-253)

```
QY 1579 GTGAATGACCAATAATAATGACTAGAGAGAAAGATGAAGATTGTCATGAATTAAG 1638
DB 1 MetLysGlyProIleIleMetThrArgGluArgMetLysIleValHisGluIleLys 20

QY 1639 GAACGAATATTGGATAAATATGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCT 1698
DB 21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40

QY 1699 GGTCTGACAGTGTATGGGCCCTATTTCGATATTGAGATGATGTGTGCATGTCAACAGAG 1758
DB 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetCysValMetSerThrGlu 60

QY 1759 GAACGAGAGTTCAGCCATGAATGACACAACCGGTGAGTGAAGGTGGAAGTGAATTTGAT 1818
DB 61 GlyAlaGluPheSerTyrGluTrpThrThrGlyGluTrpLysAlaGluValAsnPheTyr 80

QY 1819 AGCAAGAGAGTTCCTACTAGATTATGCATCTCAGGTGGAATCAGATTGGCCGCTTACACAT 1878
DB 81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluSerAspTrpProLeuThrHis 100

QY 1879 GGTCAATTTTCTCTATTTCGCGATTATGATTCAGTGGATACTTAGAGAAAGTGTAT 1938
DB 101 GlyArgPhePheSerIleLeuProIleTyrAspProGlyGlyTyrPheGluLysValTyr 120

QY 1939 CAAACTCTAAATCGGTAGAACCCCAACGGTTCACAGTGGATTTGTCCTTATCGTA 1998
DB 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140

QY 1999 GAAGAGCTGTTTGAATATGCAGGCAAAATGGCGTAATATTCGTGTCAGAGGACCGACA 2058
DB 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160

QY 2059 TTTCTACCATCTTCGACTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTGGTCTGCAT 2118
DB 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180

QY 2119 CATCGATCTGTTATACGACGAGCGCTTCGGTCTTAATCAAGCAGTGAAGCAATCAGAT 2178
DB 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp 200

QY 2179 CTTCCTTCAGGTATGACCATCTCTGCGAGTTCGTAATGTCGTCACTTCGCACTCT 2238
DB 201 LeuProSerGlyTyrAspHisLeuCysGlnLeuValMetSerGlyGlnLeuSerAspSer 220

QY 2239 GAGAACTCTTCGGAATCGCTAGAGAAATTCGGAATGGGATTCAGGAGTGGACAGAACGA 2298
DB 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240

QY 2299 CACGGATATATAGTGGATGTCTCAAAACGCATACCATTT 2337
DB 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253
```

## RESULT 9

```
US-10-718-628-13
; Sequence 13, Application US/10718628
; Publication No. US20040077016A1
; GENERAL INFORMATION:
; APPLICANT: YOKOYAMA, SIGEYUKI
```

APPLICANT: HOSEKI, JUN  
APPLICANT: YANO, TAKATO  
APPLICANT: KOYAMA, YOSHINORI  
APPLICANT: KURAMITSU, SEIKI  
APPLICANT: KAGAMIYAMA, HIROYUKI  
TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD  
TITLE OF INVENTION: OF SCREENING THERMOPHILIC BACTERIA USING THE SAME  
FILE REFERENCE: 04853.0048-00000  
CURRENT FILING DATE: 2003-11-24  
PRIOR APPLICATION NUMBER: US/10/718, 628  
PRIOR FILING DATE: 2003-11-24  
PRIOR APPLICATION NUMBER: US/09/697, 186B  
PRIOR FILING DATE: 2000-10-27  
PRIOR APPLICATION NUMBER: JP 309616/1999  
PRIOR FILING DATE: 1999-10-29  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 13  
LENGTH: 253  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme  
OTHER INFORMATION: obtained by introduction of point mutation into  
OTHER INFORMATION: wild type KNT gene of Staphylococcus aureus and  
OTHER INFORMATION: its expression  
US-10-718-628-13

Alignment Scores:  
Pred. No.: 1.42e-116 Length: 253  
Score: 1267.00 Matches: 240  
Percent Similarity: 96.4% Conservations: 4  
Best Local Similarity: 94.9% Mismatches: 9  
Query Match: 23.0% Indels: 0  
DB: 0 Gaps: 0

US-10-681-086-1 (1-3156) x US-10-718-628-13 (1-253)

QY 1579 GTGAATGGACCAATAATAATGACTAGAGAAAGAAATGAAGATTGTTCAATGAATTAAG 1638  
Db 1 MetAsnGlyProIleMetThrArgGluGluArgMetLysValHisGluLeuLys 20

QY 1639 GAACGAATATGGAATAATATGGGATGATGTTAAGCTATTGGTGTGTTATGGCTCTCTT 1698  
Db 21 GluArgileuAsnLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40

QY 1699 GGTGCTCAGACTGATGGCCCTATTTCGATATTGAGATGATGTCATGTCACACAGAG 1758  
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetCysValLeuSerThrGlu 60

QY 1759 GAAGCAGAGTTCAGCCATGAATGACAAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTAT 1818  
Db 61 GlyValGluPheSerHisGluTyrThrGlyGluTyrLysAlaGluValAsnPheTyr 80

QY 1819 AGCGAAGAGATTCTACTAGATTATGATCTCAGGTGAATCAGATTGGCCGCTTACACAT 1878  
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluProAspTyrProLeuThrHis 100

QY 1879 GGTCAATTTTCTCTATTTTCCGATTATGATTTCAGTGGATCTAGAGAAAGTGTAT 1938  
Db 101 GlyArgPhePheSerIleLeuProIleTyrAspSerGlyTyrLeuGlyLysValTyr 120

QY 1939 CAAATGCTAAATCGGTAGAACCCCAACCGTTCACGATCGGATTTGTGCGCTTATCGTA 1998  
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAlaIleCysAlaLeuVal 140

QY 1999 GAAGAGCTGTTGAATGATGAGCAATGCGTAATATTTCGTGTGCAAGCAGACACA 2058  
Db 141 GluGluLeuPheGluTyrAlaGlyLysTyrArgAsnIleArgValGlnGlyProThrThr 160

QY 2059 TTTCTACCATCTTACAGTAGCAATGCGAGTGCCTGATTTGATTGCTCTGCAT 2118  
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180

QY 2119 CATCGCATCTGTTATATACGACGCGCTTCGCTCTTAACCTGAAGCAGTAAACAATCAGAT 2178  
Db 181 HisArgileCysTyrThrThrSerAlaLeuValLeuThrGluAlaValLysGlnSerAsp 200

QY 2179 CTTCCCTTCAGGTTATGACCATCTGTCGAGTTCGTAATGTCGTCACACTTCCGACTCT 2238  
Db 201 LeuProSerGlyTyrAspHisLeuValMetSerGlyGlnLeuSerAspSer 220

QY 2239 GAGAAACTTCTCGAAATCGCTAGAGAATTTCTGAAATGGGATTTCAGGAGTGCACAGACGA 2298  
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTyrPasnGlyIleGlnGluTyrThrGluArg 240

QY 2299 CACGATATATAGTGGATGTGTCAAAACGACATACCATTT 2337  
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 10  
US-10-718-628-20  
Sequence 20, Application US/10718628  
Publication No. US20040077016A1  
GENERAL INFORMATION:  
APPLICANT: YOKOYAMA, SIGEYUKI  
APPLICANT: HOSEKI, JUN  
APPLICANT: YANO, TAKATO  
APPLICANT: KOYAMA, YOSHINORI  
APPLICANT: KURAMITSU, SEIKI  
APPLICANT: KAGAMIYAMA, HIROYUKI  
TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD  
TITLE OF INVENTION: OF SCREENING THERMOPHILIC BACTERIA USING THE SAME  
FILE REFERENCE: 04853.0048-00000  
CURRENT APPLICATION NUMBER: US/10/718, 628  
CURRENT FILING DATE: 2003-11-24  
PRIOR APPLICATION NUMBER: US/09/697, 186B  
PRIOR FILING DATE: 2000-10-27  
PRIOR APPLICATION NUMBER: JP 309616/1999  
PRIOR FILING DATE: 1999-10-29  
NUMBER OF SEQ ID NOS: 20  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 20  
LENGTH: 253  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme  
OTHER INFORMATION: obtained by introduction of point mutation into  
OTHER INFORMATION: wild type KNT gene of Staphylococcus aureus and  
OTHER INFORMATION: its expression  
US-10-718-628-20

Alignment Scores:  
Pred. No.: 2.25e-116 Length: 253  
Score: 1265.00 Matches: 240  
Percent Similarity: 96.8% Conservations: 5  
Best Local Similarity: 94.9% Mismatches: 8  
Query Match: 22.9% Indels: 0  
DB: 0 Gaps: 0

US-10-681-086-1 (1-3156) x US-10-718-628-20 (1-253)

QY 1579 GTGAATGGACCAATAATAATGACTAGAGAAAGAAATGAAGATTGTTCAATGAATTAAG 1638  
Db 1 MetAsnGlyProIleMetThrArgGluGluArgMetLysValHisGluLeuLys 20

QY 1639 GAACGAATATGGAATAATATGGGATGATGTTAAGCTATTGGTGTGTTATGGCTCTCTT 1698  
Db 21 GluArgileuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40

QY 1699 GGTGCTCAGACTGATGGCCCTATTTCGATATTGAGATGATGTCATGTCACACAGAG 1758  
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetCysValLeuSerThrGlu 60

QY 1759 GAAGCAGAGTTCAGCCATGAATGACAAACCGGTGAGTGGAGGTGGAAGTGAATTTTGTAT 1818  
Db 61 GlyValGluPheSerHisGluTyrThrGlyGluTyrLysAlaGluValAsnPheTyr 80

QY 1819 AGCGAAGAGATTCTACTAGATTATGATCTCAGGTGAATCAGATTGGCCGCTTACACAT 1878  
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluProAspTyrProLeuThrHis 100

QY 1879 GGTCAATTTTCTCTATTTTCCGATTATGATTTCAGTGGATCTAGAGAAAGTGTAT 1938  
Db 101 GlyArgPhePheSerIleLeuProIleTyrAspSerGlyTyrLeuGlyLysValTyr 120

QY 1939 CAAATGCTAAATCGGTAGAACCCCAACCGTTCACGATCGGATTTGTGCGCTTATCGTA 1998  
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAlaIleCysAlaLeuVal 140

QY 1999 GAAGAGCTGTTGAATGATGAGCAATGCGTAATATTTCGTGTGCAAGCAGACACA 2058  
Db 141 GluGluLeuPheGluTyrAlaGlyLysTyrArgAsnIleArgValGlnGlyProThrThr 160

QY 2059 TTTCTACCATCTTACAGTAGCAATGCGAGTGCCTGATTTGATTGCTCTGCAT 2118  
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180

Db 61 GluAlaGluPheSerTyr-GluTrpThrThrGlyGluTrpLysAlaGluValAsnPheTyr 80  
QY 1819 AGCGAGAGATCTACTAGATTATGCATCTCAGGTGGAAATCAGATTGGCCCTTACACAT 1878  
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluProAspTrpProLeuThrHis 100  
QY 1879 GGTCAATTTTCTCTATTGTCGCGATTATGATTCCAGGTGGATATCTTAGAGAAAGTGTAT 1938  
Db 101 GlyLysPhePheSerIleLeuProIleTyrAspProGlyGlyTyrLeuGluLysValTyr 120  
QY 1939 CAAACTGCTAAATCGGTAGAACCCAAACGTTCCACGATGCGATTGTCGCCCTTATCGTA 1998  
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140  
QY 1999 GAAGAGCTGTTGTAATATGCGAGGCAATGGCGTAATATTCGTGCAAGGACCGACAACA 2058  
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160  
QY 2059 TTCTTACCATCCTTGACTGTACAGGTAGCAATGGCAGGTGCATGTTGATTGGTCTGCAT 2118  
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180  
QY 2119 CATCGATCTGTTATACGACGAGCGCTTCGTCTTAACGAGCAAGTTPAACCAATCAGAT 2178  
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp 200  
QY 2179 CTTCCTTCAGGTTATGACCATCTGCGCCAGTTCGTAATGTCGTCAACTTCCGACTCT 2238  
Db 201 LeuProSerGlyTyrValGlnLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220  
QY 2239 GAGAACTTCCTGGAATCGCTAGAGAAATTTCTGGAAATGGGATTTCAGGAGTGGACAGAACA 2298  
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240  
QY 2299 CACGGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337  
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

## RESULT 11

US-10-718-628-17  
; Sequence 17, Application US/10718628  
; Publication No. US20040077016A1  
; GENERAL INFORMATION:  
; APPLICANT: YOKOYAMA, SIGEYUKI  
; APPLICANT: HOSEKI, JUN  
; APPLICANT: YANO, TAKATO  
; APPLICANT: KOYAMA, YOSHINORI  
; APPLICANT: KURAMITSU, SEIKI  
; APPLICANT: KAGAMIYAMA, HIROYUKI  
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD  
; FILE REFERENCE: 04853.0048-00000  
; CURRENT APPLICATION NUMBER: US/10/718,628  
; CURRENT FILING DATE: 2003-11-24  
; PRIOR APPLICATION NUMBER: US/09/697,186B  
; PRIOR FILING DATE: 2000-10-27  
; PRIOR APPLICATION NUMBER: JP 309616/1999  
; PRIOR FILING DATE: 1999-10-20  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 17  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme  
; OTHER INFORMATION: obtained by introduction of point mutation into  
; OTHER INFORMATION: wild type KMT gene of Staphylococcus aureus and  
; OTHER INFORMATION: its expression  
US-10-718-628-17

Alignment Scores:  
Pred. No.: 4.48e-116 Length: 253

Score: 1262.00 Matches: 240  
Percent Similarity: 96.4% Conservative: 4  
Best Local Similarity: 94.9% Mismatches: 9  
Query Match: 22.9% Indels: 0  
DB: 4 Gaps: 0  
US-10-681-086-1 (1-3156) x US-10-718-628-17 (1-253)  
QY 1579 GTGAATGACCAATAATAATGACTAGACAGAGAAAGATGATGTTTCATGAAATTAAG 1638  
Db 1 MetAsnGlyProIleIleMetThrArgGluLuhArgMetLysIleValHisGluIleLys 20  
QY 1639 GAACCAATATTGGATAAATATGGGATGATGTTAAGCGTATTGCTGTTATTGGCTCTCT 1698  
Db 21 GluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyrGlySerLeu 40  
QY 1699 GGTGCTCAGACTGATGGGCCCTATTCCGATATATCAGATGATGTGTGCATGTCAACAGAG 1758  
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60  
QY 1759 GAACGAGATTCACCCATGATGACCAACCGGTGAGTGGAGGTGGAAGTGAATTTTCAT 1818  
Db 61 GlyAlaGluPheSerTyrGluTrpThrThrGlyGluTrpLysAlaGluValAsnPheTyr 80  
QY 1819 AGCGAAGAGATTTCTACTAGATTATGCATCTCAGGTGGAATCAGATTGGCCCTTACACAT 1878  
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluSerAspTrpProLeuThrHis 100  
QY 1879 GGTCAATTTTCTCTATTGTCGCGATTATGATTTCAGGTGGATCTTAGAGAAAGTGTAT 1938  
Db 101 GlyArgPhePheSerIleLeuProIleTyrAspProGlyGlyTyrPheGluLysValTyr 120  
QY 1939 CAACTCTTAATCGGTAGAGCCAAACGTTCCACGATGCGATTGTCGCCCTTATCGTA 1998  
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140  
QY 1999 GAAGAGCTGTTGTAATATGCGAGGCAATGGCGTAATATTCGTGCAAGGACCGACAACA 2058  
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160  
QY 2059 TTCTTACCATCCTTTGACTGTACAGGTAGCAATGGCAGGTGCATGTTGATTGGTCTGCAT 2118  
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180  
QY 2119 CATCGATCTGTTATACGACGAGCGCTTCGTCTTAACGAGCAAGTTPAACCAATCAGAT 2178  
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp 200  
QY 2179 CTTCCTTCAGGTTATGACCATCTGCGCCAGTTCGTAATGTCGTCAACTTCCGACTCT 2238  
Db 201 LeuProSerGlyTyrValGlnLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220  
QY 2239 GAGAACTTCCTGGAATCGGTAGAGAAATTTCTGGAAATGGGATTTCAGGAGTGGACAGAACA 2298  
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240  
QY 2299 CACGGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337  
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253  
RESULT 12  
US-10-718-628-18  
; Sequence 18, Application US/10718628  
; Publication No. US20040077016A1  
; GENERAL INFORMATION:  
; APPLICANT: YOKOYAMA, SIGEYUKI  
; APPLICANT: HOSEKI, JUN  
; APPLICANT: YANO, TAKATO  
; APPLICANT: KOYAMA, YOSHINORI  
; APPLICANT: KURAMITSU, SEIKI  
; APPLICANT: KAGAMIYAMA, HIROYUKI  
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD  
; FILE REFERENCE: 04853.0048-00000  
; CURRENT APPLICATION NUMBER: US/10/718,628  
; CURRENT FILING DATE: 2003-11-24  
; PRIOR APPLICATION NUMBER: US/09/697,186B  
; PRIOR FILING DATE: 2000-10-27  
; PRIOR APPLICATION NUMBER: JP 309616/1999  
; PRIOR FILING DATE: 1999-10-20  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 17  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme  
; OTHER INFORMATION: obtained by introduction of point mutation into  
; OTHER INFORMATION: wild type KMT gene of Staphylococcus aureus and  
; OTHER INFORMATION: its expression  
US-10-718-628-17

Qy	2239	GAGAACTTCTGGAAATCGCTAGAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGA	2298
Db	221	GlulysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg	240
Qy	2299	CACGGATATATAGTGGATGTGTCAAAACGCATACCATTT	2337
Db	241	HisGlyTrpIleValAspValSerLysArgIleProPhe	253
RESULT 13			
US-10-718-628-16			
; Sequence 16, Application US/10718628			
; Publication No. US20040077016A1			
; GENERAL INFORMATION:			
; APPLICANT: YOKOYAMA, SIGEYUKI			
; APPLICANT: HOSEKI, JUN			
; APPLICANT: YANO, TAKATO			
; APPLICANT: KUYAMA, YOSHINORI			
; APPLICANT: KURAMITSU, SEIKI			
; APPLICANT: KAGAMIYAMA, HIROYUKI			
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD			
; TITLE OF INVENTION: OF SCREENING THERMOPHILIC BACTERIA USING THE SAME			
; FILE REFERENCE: 04853.0048-00000			
; CURRENT APPLICATION NUMBER: US/10/718,628			
; CURRENT FILING DATE: 2003-11-24			
; PRIOR APPLICATION NUMBER: US/09/697,186B			
; PRIOR FILING DATE: 2000-10-27			
; PRIOR APPLICATION NUMBER: JP 309616/1999			
; PRIOR FILING DATE: 1999-10-29			
; NUMBER OF SEQ ID NOS: 20			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 16			
; LENGTH: 253			
; TYPE: PRT			
; ORGANISM: Artificial Sequence			
; FEATURE:			
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme			
; OTHER INFORMATION: obtained by introduction of point mutation into			
; OTHER INFORMATION: wild type KNT gene of Staphylococcus aureus and			
; OTHER INFORMATION: its expression			
US-10-718-628-16			
Alignment Scores:			
Pred. No.: 7,09e-116 Length: 253			
Score: 1260.00 Matches: 237			
Percent Similarity: 96.0% Conservative: 6			
Best Local Similarity: 93.7% Mismatches: 10			
Query Match: 22.8% Indels: 0			
DB: 4 Gaps: 0			
US-10-681-086-1 (1-3156) x US-10-718-628-16 (1-253)			
Qy	1579	GTGAATGGACCAATAATGACTAGNAGAAAGATGAAGATTGTTTCATGAAATTAAG	1638
		:::	
Db	1	MetAsnGlyProIleIleMetThrArgGluGluArgMetIysIleValHisGluIleLys	20
Qy	1639	GAACGAATATTGATAAATATGGGATGATGTTAAGGCTATTGGTGTGTTATGGCTCTCTT	1698
Db	21	GluArgIleIleuAspLysTrpGlyAspAspValIysIleValGlySerIleu	40
Qy	1699	GGTCGTCAGACTGATGGCCCTATTCCGATATTGAGATGATGTGTGTCATGTCAACAGAG	1758
Db	41	GlyArgGlnThrAspGlyProTrpSerAspIleGluMetMetCysValMetSerThrGlu	60
Qy	1759	GAAGCAGAGTTACCCATGAATGCACACCGGTGAGTGGAGGTGGAGAGTGAATTTGAT	1818
Db	61	GluThrGluPheSerTrpGluTrpThrThrGlyGluTrpLysAlaGluValAsnPheTrp	80
Qy	1819	AGCGAGAGATTCTACTAGATTATGTCATCTCAGTGGAAATCAGATTGGCCGCTTACACAT	1878
Db	81	SerGluGluIleIleuAspTrpAlaSerArgValGluProAspTrpProLeuThrHis	100
Qy	1879	GGTCAATTTTTCCTATTFTTTTCCGCAATTTATGATTTCAGTGGTGAATCTATAGAGAAGTGTAT	1938

Db 101 GlyLysPheSerIleLeuProIleThrGlyGlyTyrLeuGluLysValTyr 120  
Qy 1939 CAACTGCTAAATCGGTAGAGCCCAACGGTTCACGATGGATTTGGCCCTTATCGTA 1998  
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140  
Qy 1999 GAAGAGCTGTTGTAATGTCAGGCAAAATGGCGTAAATTCGTGTGTCAGGAGCCGACAAACA 2058  
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgenIleArgValGlnGlyProLeuThr 160  
Qy 2059 TTTCTACCATCTTGTACGTAGCATAGCATGGCGTTCCTTAAGCAAGCAATTCAGAT 2118  
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180  
Qy 2119 CATGCACTCTGTTATACAGAGCGCTTCGGTCTTAAGCAAGCAATTCAGAT 2178  
Db 181 HisArgIleCysTyrThrGlyAlaSerValLeuThrGluAlaValArgGlnProAsp 200  
Qy 2179 CTTCCTTCAGGTATGACCATCTGCGCAGTTCGTAATGCTGTGTCAGTTCAGACTCT 2238  
Db 201 LeuProGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220  
Qy 2239 GAGAACTCTCGAATCCCTAGAGCAATTCGGAATGGGATTCAGGAGTGGACAGCAACA 2298  
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpAlaGluArg 240  
Qy 2299 CACGATATATAGTGATGTGTCAAAACGCATACCATTT 2337  
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

## RESULT 14

US-10-718-628-19  
; Sequence 19, Application US/10718628  
; Publication No. US20040077016A1  
; GENERAL INFORMATION:  
; APPLICANT: YOKOYAMA, SIGEYUKI  
; APPLICANT: HOSEKI, JUN  
; APPLICANT: KANO, TAKATO  
; APPLICANT: KOYAMA, YOSHINORI  
; APPLICANT: KURAMITSU, SEIKI  
; APPLICANT: KAGAMIYAMA, HIROYUKI  
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD  
; FILE REFERENCE: 04853.0048-00000  
; CURRENT APPLICATION NUMBER: US/10/718.628  
; CURRENT FILING DATE: 2003-11-24  
; PRIOR FILING DATE: 2000-10-27  
; PRIOR APPLICATION NUMBER: JP 309616/1999  
; PRIOR FILING DATE: 1999-10-29  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 19  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme  
; OTHER INFORMATION: obtained by introduction of point mutation into  
; OTHER INFORMATION: wild type Knt gene of Staphylococcus aureus and  
; OTHER INFORMATION: its expression  
US-10-718-628-19

Alignment Scores:  
Pred. No.: 1,786-115 Length: 253  
Score: 1256.00 Matches: 239  
Percent Similarity: 96.0% Conservative: 4  
Best Local Similarity: 94.5% Mismatches: 10  
Query Match: 22.8% Indels: 0  
DB: 4 Gaps: 0

US-10-681-086-1 (1-3156) x US-10-718-628-19 (1-253)

Qy 1579 GTGAATGACCAATAATAATGACTAGAGAAAGAAATGAAGATTGTTTCATGAATTAAG 1638  
Db 1 MetAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluLys 20  
Qy 1639 GAACGAATATTGGATAAAATATGGGATGATGTTTAAGGCTATTGCTGTTATGGCTCTCTT 1698  
Db 21 GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu 40  
Qy 1699 GGTGCTCAGACTGATGGCCCTATTTCGATATTGAGATGATGCTGTCATGTCACAGAG 1758  
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60  
Qy 1759 GAACACAGATTTCACCCATGAATGACAAACCGTCAGTGGAGGTGGAAGTCAATTTTCAT 1818  
Db 61 GlyAlaGluPheSerTyrGluTrpThrGlyGluTrpLysAlaGluValAsnPheTyr 80  
Qy 1819 AGCGAAGAGATTCTACTAGATTATGCATCTCAGGTGGAATCAGATTGGCCCTTACACAT 1878  
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerArgValGluProAspTrpProLeuThrHis 100  
Qy 1879 GGTCAATTTTCTCTATTTCGCGATTTATGATTCAGGTGGATACCTTAGAGAAAGTGTAT 1938  
Db 101 GlyLysPhePheSerIleLeuProIleTyrAspProGlyGlyTyrLeuGluLysValTyr 120  
Qy 1939 CAAACTGCTAAATCGGTAGAGCCCAACGGTTCACGATGCGATTGTCGCCCTTATCGTA 1998  
Db 121 GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal 140  
Qy 1999 GAAGAGCTGTTGTAATATGACGCAAAATGGCGTAAATTCGTGTGCAAGGACCGACAACA 2058  
Db 141 GluGluLeuPheGluTyrAlaGlyLysTrpArgenIleArgValGlnGlyProThrThr 160  
Qy 2059 TTTCTACCATCTTGTACGTAGCATAGCATGGCGTTCAGTTCGATTTGCTGCTGCTCAT 2118  
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180  
Qy 2119 CATGCACTCTGTTATACAGAGCGCTTCGGTCTTAAGCAAGCAATTCAGAT 2178  
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp 200  
Qy 2179 CTTCCTTCAGGTTATACCATCTGTCGCGATTCGTAATGCTGTCATCTGCTCAACTTTCGACTCT 2238  
Db 201 LeuProSerGlyTyrValGlnLeuCysGlnPheValMetSerGlyGlnLeuSerAspPro 220  
Qy 2239 GAGAACTCTCGAATCGCTAGAGCAATTCGGAATGGGATTCAGGAGTGGACAGCAACA 2298  
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240  
Qy 2299 CACGATATATAGTGATGTGTCAAAACGCATACCATTT 2337  
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

## RESULT 15

US-10-718-628-3  
; Sequence 3, Application US/10718628  
; Publication No. US20040077016A1  
; GENERAL INFORMATION:  
; APPLICANT: YOKOYAMA, SIGEYUKI  
; APPLICANT: HOSEKI, JUN  
; APPLICANT: YANO, TAKATO  
; APPLICANT: KOYAMA, YOSHINORI  
; APPLICANT: KURAMITSU, SEIKI  
; APPLICANT: KAGAMIYAMA, HIROYUKI  
; TITLE OF INVENTION: MUTANT KANAMYCIN NUCLEOTIDYLTRANSFERASE AND A METHOD  
; FILE REFERENCE: 04853.0048-00000  
; CURRENT APPLICATION NUMBER: US/10/718.628  
; CURRENT FILING DATE: 2003-11-24  
; PRIOR APPLICATION NUMBER: US/09/697,186B  
; PRIOR FILING DATE: 2000-10-27  
; PRIOR APPLICATION NUMBER: JP 309616/1999  
; PRIOR FILING DATE: 1999-10-29

; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 253  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Mutant enzyme  
; OTHER INFORMATION: obtained by introduction of point mutation into  
; OTHER INFORMATION: wild type KNT gene of Staphylococcus aureus and  
; OTHER INFORMATION: its expression

US-10-718-628-3

Alignment Scores:  
Pred. No.: 4.33e-112 Length: 253  
Score: 1222.00 Matches: 231  
Percent Similarity: 93.7% Conservativeness: 6  
Best Local Similarity: 91.3% Mismatches: 16  
Query Match: 22.1% Indels: 0  
DB: 4 Gaps: 0

US-10-681-086-1 (1-3156) x US-10-718-628-3 (1-253)

Qy	1579	GTGAATCGACATATAATGACTAGAGAGAAAGATGTTTCATGAAATTAAG	1638
Db	1	MetLysGlyProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys	20
Qy	1639	GAACGATATTGGTAATATGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCTT	1698
Db	21	GluArgIleLeuAspLysTyrGlyAspValLysAlaIleGlyValTyrGlySerLeu	40
Qy	1699	GGTCGTGACAGTCAATGGCCCTATTCCGATATTGAGATGATGTGTCAACAGAG	1758
Db	41	GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValLeuSerThrGlu	60
Qy	1759	GAAGCAGATTGACCCATGAATGAGCAACCCGGTGGAGTGGAAAGTGAATTTTGAT	1818
Db	61	GlyValGluPheSerTyrGluTyrThrGlyGluTyrLysAlaGluValAsnPheTyr	80
Qy	1819	AGCGAGAGATTCCTACTAGATTATGTCATCTCAGTGGATTCAGATTCGCCCTTACACAT	1878
Db	81	SerGluGluIleLeuLeuAspTyrAlaSerArgValGluProAspTrpProLeuThrHis	100
Qy	1879	GGTCAATTTTCTCTATTTTCCGATTTATGATTCAAGTGGATCTTAGAGAAAGTGTAT	1938
Db	101	GlyArgPhePheSerIleLeuProIleTyrAspProGlyGlyTyrPheGluLysValTyr	120
Qy	1939	CAAACTGCTAAATCGGTAGAACCCAAACGTTCCACGATGCGATTTGTGCCCTTATCGTA	1998
Db	121	GlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAlaLeuIleVal	140
Qy	1999	GNAGAGCTGTTGAATATGAGGCAATGGGTAAATTCGTGTGCAAGGACCGACAACA	2058
Db	141	GluGluLeuPheGluTyrAlaGlyLysTyrArgAsnIleArgValGlnGlyProThrThr	160
Qy	2059	TTTCTACATCTTGACTGTACAGGTAGCAATGGCAGTCCATGTGATTGGTCTGCAT	2118
Db	161	PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis	180
Qy	2119	CATCGCATCTGTTATACAGAGCGCTTCGGTCTTAACTGAAGCAGTTTAAGCAATCAGAT	2178
Db	181	HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnProAsp	200
Qy	2179	CTTCTCTCAGGTTATGACCATCTGTGCCAGTTCGTAAATGCTGGTCAACTTCCGACTCT	2238
Db	201	LeuProProGlyTyrValGlnLeuCysGlnLeuValMetSerGlyGlnLeuSerAspPro	220
Qy	2239	GAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTACAGGAGTGACAGAACGA	2298
Db	221	GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyValGlnGluTyrAlaGluArg	240
Qy	2299	CACGGATATATAGTGGATGTGTCAAAACGCGATACCATTT	2337

Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253  
Search completed: February 10, 2006, 04:06:10  
Job time : 456.5 secs

***This Page Blank (uspto)***



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - protein search, using frame\_plus\_n2p model

Run on: February 10, 2006, 03:51:42 ; Search time 4 Seconds  
(without alignments)

2070.736 Million cell updates/sec

Title: US-10-681-086-1

Perfect score: 5517

Sequence: 1 ggatccacgaggttagcagc.....atcggtatgatgtcgatttc 3156

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Xgapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 97014 seqs, 13122538 residues

Total number of hits satisfying chosen parameters: 194028

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlp

-Q=/abse/ABSSWEB\_spool/US10681086/runat\_09022006\_140529\_12034/app\_query.fasta.1

-DB=Published Applications AA New -QFMT=fastan -SUFFIX=n2p.rapbn -MINMATCH=0.1

-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62

-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100

-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZ=500 -MINLEN=0

-MAXLEN=200000000 -HOST=abs02p

-USER=US10681086 @CGN\_1\_1\_7 @runat\_09022006\_140529\_12034 -NCPU=6 -ICPU=3

-NO MMAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV TIMEOUT=120

-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7

-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications AA New:

1: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.ppe.\*

2: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.ppe.\*

3: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.ppe.\*

4: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.ppe.\*

5: /cgn2\_6/ptodata/1/pubppaa/US05\_NEW\_PUB.ppe.\*

6: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.ppe.\*

7: /cgn2\_6/ptodata/1/pubppaa/US11\_NEW\_PUB.ppe.\*

8: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.ppe.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	828	15.0	332	6	US-10-793-626-540
2	828	15.0	332	6	US-10-793-626-2562
3	548	9.9	381	7	US-11-098-686-10666
4	545.5	9.9	358	6	US-10-467-657-6970
5	405.5	7.4	322	7	US-11-098-686-10665
6	403	7.3	377	6	US-10-793-626-2926
7	373.5	6.8	334	7	US-11-055-822-644
8	350	6.3	416	6	US-10-467-657-5060
9	115	2.1	339	6	US-10-467-657-4318

10	111.5	2.0	401	6	US-10-467-657-1860	Sequence 1860, Ap
11	110	2.0	588	7	US-11-186-284-213	Sequence 213, App
12	110	2.0	596	6	US-10-821-234-1068	Sequence 1068, Ap
13	104	1.9	425	6	US-10-525-710-22	Sequence 22, Appl
14	103.5	1.9	5935	6	US-10-595-561-776	Sequence 776, App
15	102	1.8	348	7	US-11-055-822-662	Sequence 662, App
16	102	1.8	426	6	US-10-525-710-28	Sequence 28, Appl
17	100.5	1.8	464	7	US-11-057-012-41	Sequence 41, Appl
18	98	1.8	4655	6	US-10-595-561-556	Sequence 556, App
19	96.5	1.7	391	7	US-11-055-822-60	Sequence 60, Appl
20	96.5	1.7	391	7	US-11-055-822-300	Sequence 300, App
21	96.5	1.7	391	7	US-11-055-822-328	Sequence 328, App
22	96	1.7	386	6	US-10-467-657-6564	Sequence 6564, Ap
23	95.5	1.7	417	6	US-10-858-730-227	Sequence 227, App
24	95	1.7	950	7	US-11-077-550-118	Sequence 118, App
25	94.5	1.7	428	6	US-10-525-710-20	Sequence 20, Appl
26	94.5	1.7	430	6	US-10-525-710-30	Sequence 30, Appl
27	94	1.7	1216	7	US-11-115-639-44	Sequence 44, Appl
28	94	1.7	1216	7	US-11-115-639-45	Sequence 45, Appl
29	92.5	1.7	423	6	US-10-525-710-44	Sequence 44, Appl
30	92.5	1.7	5024	6	US-10-793-626-2964	Sequence 2964, Ap
31	92	1.7	437	6	US-10-858-730-214	Sequence 214, App
32	92	1.7	437	6	US-10-858-730-283	Sequence 283, App
33	91.5	1.7	4544	7	US-11-076-427A-32	Sequence 32, Appl
34	91	1.6	423	7	US-11-055-822-640	Sequence 640, App
35	90.5	1.6	338	6	US-10-467-657-4976	Sequence 4976, Ap
36	90.5	1.6	345	7	US-11-124-368A-285	Sequence 285, App
37	90.5	1.6	514	6	US-10-840-688-8	Sequence 8, Appl
38	90.5	1.6	907	7	US-11-124-368A-288	Sequence 288, App
39	89.5	1.6	805	6	US-10-467-657-8246	Sequence 8246, Ap
40	89.5	1.6	912	7	US-11-077-550-116	Sequence 116, App
41	89.5	1.6	1706	7	US-11-052-554A-176	Sequence 176, App
42	89	1.6	437	6	US-10-858-730-288	Sequence 288, App
43	89	1.6	914	7	US-11-077-550-120	Sequence 120, App
44	88.5	1.6	944	7	US-11-077-550-122	Sequence 122, App
45	88.5	1.6	970	6	US-10-467-657-5524	Sequence 5524, Ap

#### ALIGNMENTS

##### RESULT 1

US-10-793-626-540

; Sequence 540, Application US/10793626

; Publication No. US20050255478A1

; GENERAL INFORMATION:

; APPLICANT: KIMMERLY WILLIAM JOHN

; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS

; FILE REFERENCE: P034800S

; CURRENT APPLICATION NUMBER: US/10/793,626

; CURRENT FILING DATE: 2004-03-04

; PRIOR APPLICATION NUMBER: 60/164,258

; PRIOR FILING DATE: 1999-11-09

; NUMBER OF SEQ ID NOS: 4472

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 540

; LENGTH: 332

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: synthetic

; OTHER INFORMATION: amino acid sequence

US-10-793-626-540

Alignment Scores:

Pred. No.: 2.49e-67 Length: 332

Score: 828.00 Matches: 158

Percent Similarity: 75.0% Conservative: 49

Best Local Similarity: 57.2% Mismatches: 51

Query Match: 15.0% Indels: 18

DB: 6 Gaps: 2

US-10-681-086-1 (1-3156) x US-10-793-626-540 (1-332)

Qy	2307	TATAGTCGATGTGTC	CAAAACGCATACCATTTTG	AAATTCGAAAGCGCGGATTT	-----	2357
Db	57	TyGlyLyValysValys	LeuAsnMetIleLeuAsnAla	lysSerGlyIleCysAlaGlu	76	
Qy	2358	-----	-----	-GAGTCTTACCGGATGGTG	2375	
Db	77	AspCysGlyTyrCys	GlyGlnSerValysMet	LysGlyGlnArgTyrAlaLeuVal	96	
Qy	2376	AATAAGGAAACGCTG	CTTGAAGCGCGAAGCGCGC	ACGACGATCTGTAATATCGGCACATAT	2435	
Db	97	GluGlnAspGlnIle	LysGlyAlaGlnValAla	ThrGluAsnGlnIleGlyThrTyr	116	
Qy	2436	TGTATCTGTCGACG	CGCAGAGGTCCGTCTAA	CACGAGAACTGGATCAGGTCTGATAGTCG	2495	
Db	117	CysIleValMetSer	GlyArgGlyProSer	AsnArgGluValAspHisIleCysGlyThr	136	
Qy	2496	GTTCAGGAAATTAAG	AGACGTAT--GGACTGA	AGATTTGTGCATGCTCTTGACACTGTTG	2552	
Db	137	ValGluAspIleLe	LysIleHisProGlnLeu	lysIleCysAlaCysLeuGlyLeuThr	156	
Qy	2553	AAGCCAGACGAGCG	CAAGCGCTCAAGATG	CGAGGTAGCACCGCTATAATCATATTTG	2612	
Db	157	LysGluGluGlnAla	LysLeuLysAlaAla	GlyValAspArgTyrAsnHisAsnLeu	176	
Qy	2613	AATACGTCACAGAA	ACCATTCAAACATCAC	AACTCACATACATACGATCAGCAGTCC	2672	
Db	177	AsnThrSerGluArg	TyrHisAspGluVal	ValThrHisThrTyrGluAspArgVal	196	
Qy	2673	AATACGGTTCGAAAT	CGCAAAAGAAATCGGG	CGTGTCTCCGTGTTCAGCGCGCATATTCGGG	2732	
Db	197	AsnThrValGluMet	LysAspAsnAsnIle	SerProCysSerGlyValIleCysGly	216	
Qy	2733	ATGAGGAGACGAGAA	CAGGATGTCATTGCAT	CGCCAAAGCTTGAAGGCTCTTGACGGC	2792	
Db	217	MetGlyGluSerAsn	GlnAspIleIleAspMet	AlaPheAlaLeuArgAlaIleAspAla	236	
Qy	2793	GATTCCATTCTGTGA	ATTTTTCATGCAATTG	TGGCAGCGCGTTAGAAAGCGCTCAAC	2852	
Db	237	AspSerIleProIle	AsnPheLeuHisPro	IleLysGlyThrLysPheGlyGlyLeuAsp	256	
Qy	2853	GAATTAAACCGCTG	TATGTTTAAAGTGCTG	GGCGTGTTCGGTTTTATCAATCCATCA	2912	
Db	257	LeuLeuSerProMet	LysCysLeuArgIle	IleAlaMetPheArgLeuIleAsnProThr	276	
Qy	2913	AAAGAAATTCGCAT	TTCCGGAGNAGAGAGGT	CAATCTCCGACATTCGACCATTAGGG	2972	
Db	277	LysGluIleArgIle	AlaGlyArgGluVal	AsnLeuArgSerLeuGlnProLeuAla	296	
Qy	2973	CTTTACCGCCCAACT	CCATTTTGTCCGAGACT	TACTTAACAACCTCGCGCAAGAGGAG	3032	
Db	297	LeuLysAlaAla	AsnSerIlePheVal	GlyAspTyrLeuIleThrGlyGlyGlnProAsn	316	
Qy	3033	ACGAGAGATCATNA	AAATCTGCTAGTGAT	TTTAGCTTTGAGCTGAATCA	3080	
Db	317	GluGluAspTyrArg	MetIleGluAspLeu	GlyPheGluIleAspSer	332	

## RESULT 2

US-10-793-626-2562  
 ; Sequence 2562, Application US/10793626  
 ; Publication No. US20050255478A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: KIMMERLY, WILLIAM JOHN  
 ; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
 ; FILE REFERENCE: PU3480US  
 ; CURRENT APPLICATION NUMBER: US/10/793,626  
 ; CURRENT FILING DATE: 2004-03-04  
 ; PRIOR APPLICATION NUMBER: 60/164,358  
 ; PRIOR FILING DATE: 1999-11-09  
 ; NUMBER OF SEQ ID NOS: 4472  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 2562  
 ; LENGTH: 332  
 ; TYPE: PRT

```
US-11-098-686-10666
; Sequence 10666, Application US/11098686
; Publication No. US2006024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10666
; LENGTH: 381
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10666

Alignment Scores:
Pred. No.: 7.32e-42 Length: 381
Score: 548.00 Matches: 131
Percent Similarity: 56.9% Conservative: 54
Best Local Similarity: 40.3% Mismatches: 132
Query Match: 9.9% Indels: 8
DB: 7 Gaps: 3

US-10-681-086-1 (1-3156) x US-11-098-686-10666 (1-381)
QY 151 TGGTCTCTCAAAATATTAGGCTCGCAAGCATAGAGCTTTGATCGATGAGCCCAA 210
Db 35 PheSerThrAsnAspTyrLeuGlnLeuAlaTyrHisProAsnLeuLeuAsnAlaIle 54
QY 211 ACAGCATTCGACCAATTTGGGACAGGAGCGGTTTACGTTTAAACGACAGGCAATTCG 270
Db 55 MetThrGlySerThrTyrGlyValGlySerThrGlySerArgLeuLeuSerGlyAsnAsn 74
QY 271 GTCTGGCATGAAAGCTAGAAAAGAGATTGCCAGCTTTAACTGACAGAGCGGCTCG 330
Db 75 GluLeuPheGluArgLeuGluThrThrIleAlaGlnAspLysHisThrGluThrAlaMet 94
QY 331 CTGTTTTCGAGCGGTACTTGGCCAATGTCGGTCTCTTTCATCTTGCCTGCGCAGAA----- 384
Db 95 LeuPheIleSerGlyPheGlnThrAsnValSerValLeuSerAlaLeuLeuAspHisHis 114
QY 385 -----AAGGAAGATGTCAATTTAAGTGACCGCTCAATCATGCAAGTATGATCGAC 435
Db 115 ValLeuLysMetGlnProLeuValPhePheAspLysLeuAsnHisAlaSerLeuTyrGln 134
QY 436 GGCTGCGCATTTCTAAGGCTGTACAGTTGTTTATCGGCATATTTGATGATGATCTTT 495
Db 135 AlaValPheLeuSerLysAlaGluLeuLeuArgTyrTyrHisAsnAsnMetGluHisLeu 154
QY 496 GAAACAACAGCTGAATGAACACACAGCGTTATCAGCGC---CGTTTATCGTAAACAGCGGA 552
Db 155 SerSerLeuLeuLysLysTyrLysAspAsnArgProLysPheIleValThrGluThr 174
QY 553 GTATTTCAGCATGGATGGCAACATCGCCCTCTTTCATCAGATCATCTCATCTTGGCAACGC 612
Db 175 LeuPheGlyMetAspGlyAspIleAlaProLeuThrAspIleValSerLeuSerSerGln 194
QY 613 TATCATCCCTTCGTGTCGTGATGATGCCAGCAACAGAGATTTTGGCGGATTCGGGA 672
Db 195 TyrGlnThrPheLeuTyrLeuAspGluAlaHisAlaThrGlyLeuThrGlyIleHisGly 214
QY 673 CAAGGACAGAGTGAATACTTTGGTGT-----TGTCGCCACATTTGTTATCGGCACCTTA 726
Db 215 TyrGlyLeuSerThrThrValAsnLeuSerHisIleProHisIleValMetGlyThrPhe 234
QY 727 AGCAAGAGCTGTGGCGCGGAGGAGGTTTTTGGCGCAGGATCAGCGGCTCTTCATCGACTTT 786

US-10-467-657-6970
; Sequence 6970, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SPA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASIGNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 6970
; LENGTH: 358
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-6970

Alignment Scores:
Pred. No.: 1.2e-41 Length: 358
Score: 545.50 Matches: 133
Percent Similarity: 58.0% Conservative: 63
Best Local Similarity: 39.3% Mismatches: 127
Query Match: 9.9% Indels: 15
DB: 6 Gaps: 7

US-10-681-086-1 (1-3156) x US-10-467-657-6970 (1-358)
QY 121 AGGAATATTGATGGCGAAAATCAACGGTC-----TGGTCTCTCAAAATATTATTAGG 174
Db 6 ArgTyrIleThrArgGluAsnArgLysMetLeuAsnMetSerSerAsnAspTyrLeuGly 25
QY 175 CTCGCAAGCGATAGAGCTTTGATCGATCGAGCCCAACAGCATTCGACGAATTTGGACA 234
Db 26 LeuAlaSerAspGluAsnLeu-----ArgArgSerPheLeuGlnGlnTyrGly 42
QY 235 GGA-----AGCAGCGGTTTACGTTTAAACGACGCAATTCGGTCTGGCAT 279
Db 43 AsnPheProSerPheThrSerSerArgLeuLeuThrGlyAsnPheProIleTyr 62
```

```
QY 280 GAAAGCTAGAAAGAGATTGCCAGCTTTAAACTGACAGAAAGCGCCCTGCTGTTTCG 339
Db      |||||:::  ::|||  |||||::|
QY 63 ThrAspLeuGluLeuValAlaGlnArgPheGlnArgGluSerAlaLeuLeuPheAsn 82
Db      |||||:::  ::|||  |||||::|
QY 340 AGCGTTTACTTGGCAATGTCGGTGTCTTTCATCTTCCTTGGCCAGAAAGAGATGTCATT 399
Db      |||||:::  ::|||  |||||::|
QY 83 SerGlyTyrHisAlaAsnLeuGlyLeuProAlaLeuThrThrLysSerLeuLeu 102
Db      |||||:::  ::|||  |||||::|
QY 400 TTAAGTGACCAAGCTCAATCATGCGATATGATGACCGCTCCGACACTTCTTAAGGTCAT 459
Db      |||||:::  ::|||  |||||::|
QY 103 LeuAlaAspLysPheValHisAlaSerMetIleAspGlyIleArgLeuSerArgCysAla 122
Db      |||||:::  ::|||  |||||::|
QY 460 ACAGTTGTTTATCGCATATTGATGATGAATGATCTTGAACCAAGCTGAATGAA---ACA 516
Db      |||||:::  ::|||  |||||::|
QY 123 PhePheArgTyrArgHisAsnAspTyrGluHisLeuLysAsnLeuLeuGluLysAsnVal 142
Db      |||||:::  ::|||  |||||::|
QY 517 CAGCGTTATAGCCCGCTTTTATCGTAACAGACGGAGTATTACGATGGATGGCACAAATC 576
Db      |||||:::  ::|||  |||||::|
QY 143 GlyLysPheAspArgThrPheIleValThrGluSerValPheSerMetAspGlyAspVal 162
Db      |||||:::  ::|||  |||||::|
QY 577 GCCCTCTTGATCAGATCATCTCACTTCGGAACGCTAT---CATGCCCTTCGTGGTCTT 633
Db      |||||:::  ::|||  |||||::|
QY 163 AlaAspLeuLysGlnLeuValGlnLeuLysGlnPheProAsnThrTyrLeuTyrVal 182
Db      |||||:::  ::|||  |||||::|
QY 634 GATGATGCCACGCAACAGGAGTTTGGCGGATTCGGGACAAAGCAAGCAGTGAATACTTT 693
Db      |||||:::  ::|||  |||||::|
QY 183 AspGluAlaHisAlaIleGlyValTyrGlyGlnAsnGlyLeuGlyIleAlaGluArgAsp 202
Db      |||||:::  ::|||  |||||::|
QY 694 GGTGTTTGTCCC-----GACATTGTTATCGCACCTTTAAGCAAAAGCTGTTGGCCGGAA 747
Db      |||||:::  ::|||  |||||::|
QY 203 AsnLeuIleAlaGluIleAspLeuLeuValGlyThrPheGlyLysAlaLeuAlaSerVal 222
Db      |||||:::  ::|||  |||||::|
QY 748 GGAGTTTGGCGGAGGATCAGCGGTCTTCATCAGCTTTTTCGTGAACCATCGCCAGACA 807
Db      |||||:::  ::|||  |||||::|
QY 223 GlyAlaTyrAlaValCysAsnGlnValLeuLysGluCysLeuIleAsnGlnMetArgPro 242
Db      |||||:::  ::|||  |||||::|
QY 808 TTTATCTTTCAAAACCGCTATTCCGCGACCGCAGCTGTGCGGTGCTCAGCGCTTTCAAC 867
Db      |||||:::  ::|||  |||||::|
QY 243 LeuIlePheSerThrAlaLeuProPheAsnValAlaTyrThrTyrPheIlePheGlu 262
Db      |||||:::  ::|||  |||||::|
QY 868 ATCATTTGAAGCCAGCAGGGGAAACAGCAGCTTTTATTTTTCATATCATCATGATCAGA 927
Db      |||||:::  ::|||  |||||::|
QY 263 ArgLeuProGlnPheSerLysGluArgSerHisLeuGluGlnLeuSerAlaPheLeuArg 282
Db      |||||:::  ::|||  |||||::|
QY 928 ACCAGCTCGAAGATATCGGTATGTGTGTGAAGAGATCACACCGCATTTTCTCTGA 987
Db      |||||:::  ::|||  |||||::|
QY 283 ArgGluValAlaHisArgThrGlnIleMetProSerGlu---ThrCysIleValProTyr 301
Db      |||||:::  ::|||  |||||::|
QY 988 GTCATTGGCGATGCCCATAAACCGTCTCTATTGCTGAAAAACTGCGGGCAAGGGAATT 1047
Db      |||||:::  ::|||  |||||::|
QY 302 IleLeuGlyGlyAsnGluAlaThrLeuAlaLysAlaGluTyrLeuGlnGlyGlnGlyTyr 321
Db      |||||:::  ::|||  |||||::|
QY 1048 TATGCTCTGCCATTCCGCCCCCAACCGTTCCGCCGGGTGAAGCCGATTCGA 1101
Db      |||||:::  ::|||  |||||::|
QY 322 TyrCysLeuProIleGlyProProThrValProLysAsnThrSerArgIleArg 339
Db      |||||:::  ::|||  |||||::|
```

## RESULT 5

```
US-11-098-686-10665
; Sequence 10665, Application US/11098686
; Publication No. US20060024696A1
; GENERAL INFORMATION:
; APPLICANT: Kapur, Vivek and Gebhart, Connie J.
; TITLE OF INVENTION: NUCLEIC ACID AND POLYPEPTIDE SEQUENCES
; FILE OF INVENTION: FROM LAWSONIA INTRACELLULARIS AND METHODS OF USING
; FILE REFERENCE: 09531-128001
; CURRENT APPLICATION NUMBER: US/11/098,686
; CURRENT FILING DATE: 2005-04-04
; PRIOR APPLICATION NUMBER: PCT/US03/31318
; PRIOR FILING DATE: 2003-10-01
; PRIOR APPLICATION NUMBER: US 60/416,395
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 11433
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 10665
; LENGTH: 322
; TYPE: PRT
; ORGANISM: Lawsonia intracellularis
US-11-098-686-10665
Alignment Scores:
Pred. No.: 6,08e-29 Length: 322
Score: 405.50 Matches: 90
Percent Similarity: 58.9% Conserv: 56
Best Local Similarity: 36.3% Mismatches: 93
Query Match: 7.4% Indels: 9
DB: 7 Gaps: 5
```

US-10-681-086-1 (1-3156) x US-11-098-686-10665 (1-322)

```
QY 2346 AAAGCGCCGATGAGTCTTACCGGATGGTGAATAAGAAACGCTGCTTGAAGCGCGAAG 2405
Db      |||||:::  ::|||  |||||::|
QY 71 LysThrAspLeuGlnLysGluProLeuValAspIleAspThrLeuIleSerGluAlaLys 90
Db      |||||:::  ::|||  |||||::|
QY 2406 CGGGCGCACGATCTGAATATCGGCACATATTGTATCTGTCGCAAGCGCGCAGAGGTCCTCT 2465
Db      |||||:::  ::|||  |||||::|
QY 91 LysAlaLysGluLeuGlySerThrArgPheCysMetGlyAlaAlaTyrAspSerProLeu 110
Db      |||||:::  ::|||  |||||::|
QY 2466 AACAGAGAAGTGGATCAGGTCTAGATGCGGTTTCAGGAAATTTAAAGAGACGTATGGACTG 2525
Db      |||||:::  ::|||  |||||::|
QY 111 AspLysAspLeuLysIleValCysGlnMetIleGluGluValLysLys---LeuGlyLeu 129
Db      |||||:::  ::|||  |||||::|
QY 2526 AAGATTGTCATGTCCTTGACACTGTTGAAGCCAGAGCAGCGGAAAGCGCTCAAGATGCA 2585
Db      |||||:::  ::|||  |||||::|
QY 130 GluThrCysValThrLeuGlyPheLeuLysGluHisGlnIleAlaMetLeuLysLysAla 149
Db      |||||:::  ::|||  |||||::|
QY 2586 GGAGTAGACCGCTATATCATTAATACGTTCACAGAGAACACCATCTCAACATCACA 2645
Db      |||||:::  ::|||  |||||::|
QY 150 GlyLeuAspPheTyrAsnHisAsnMetAsnThrSerProGluPheTyrGluHisIleAla 169
Db      |||||:::  ::|||  |||||::|
QY 2646 ACCTCACATACATACGATGACAGATCAATACGCTTCAAAATCGCAAAAGAAATCGGGCTG 2705
Db      |||||:::  ::|||  |||||::|
QY 170 ThrThrHisThrPheAspAspArgLeuAlaThrLeuLysAlaValArgLysPheGlyIle 189
Db      |||||:::  ::|||  |||||::|
QY 2706 TCTCCGTTTCAGCGCCATTTATCGGGATGAAGAGACGAGAACAGAGATGTCATTGACATC 2765
Db      |||||:::  ::|||  |||||::|
QY 190 LysLeuCysSerGlyGlyIleIleGlyLeuGlyGluThrIleAspAspArgIleSerMet 209
Db      |||||:::  ::|||  |||||::|
QY 2766 GCCAAAAGCTTGAAGGCTCTTGAC-----GCGATTTCCTCTCTGTGAATTTTTCAT 2819
Db      |||||:::  ::|||  |||||::|
QY 210 LeuLeuLeuLeuAlaThrLeuGluGluGlnProGluSerValProIleAsnArgPheVal 229
Db      |||||:::  ::|||  |||||::|
QY 2820 GCAATTGATGGCACCGCGTTAGAAAGCGGTCAACGAATTAACCCGCTGTATTGTTTAAAA 2879
Db      |||||:::  ::|||  |||||::|
QY 230 LysValAlaGlyThrProLeuAsnProGlnSerAspIleAspProPheAspPheValArg 249
Db      |||||:::  ::|||  |||||::|
QY 2880 GTGCTGCGCTGTTCCGTTTATCAATCCATCAAAAGAAATTCGCATTTCGCGAGGAAGA 2939
Db      |||||:::  ::|||  |||||::|
QY 250 IleIleAlaLeuThrArgIleLeuMetProLysSerTyrIleArgLeuAlaAlaGlyArg 269
Db      |||||:::  ::|||  |||||::|
QY 2940 GAGTCT-----AATCTCCGACATTCGCCCATTCAGGCTTTACGCCGCAAACTCC 2990
Db      |||||:::  ::|||  |||||::|
QY 270 GluGlnMetSerAspGluLeuGlnThrLeuCysPheMetGly-----GlyAlaAsnSer 287
Db      |||||:::  ::|||  |||||::|
QY 2991 ATTTTTCGCGAGACAC---TACTTTAAACAACTGCCGGCAAGAGAGACGCGAGGATCATAAA 3047
Db      |||||:::  ::|||  |||||::|
QY 288 IlePheTyrGlyArgLeuLeuThrThrAspGlyProGlnProGluGlnAspLeu 307
Db      |||||:::  ::|||  |||||::|
QY 3048 ATGCTGAGTGATTAGGCTTTGAA 3071
Db      |||||:::  ::|||  |||||::|
QY 308 LeuPheGlnLysIleGlyLeuGlu 315
Db      |||||:::  ::|||  |||||::|
```

## RESULT 6

```
US-10-793-626-2926
; Sequence 2926, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
```

```
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PUS480US
; CURRENT APPLICATION NUMBER: US/10/793,626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2926
; LENGTH: 377
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
US-10-793-626-2926

Alignment Scores:
Pred. No.: 1,09e-28 Length: 377
Score: 403.00 Matches: 114
Percent Similarity: 51.1% Conservative: 76
Best Local Similarity: 30.6% Mismatches: 138
Query Match: 7.3% Indels: 44
DB: 6 Gaps: 10

US-10-681-086-1 (1-3156) x US-10-793-626-2926 (1-377)

Qy 40 TTAACGCGGTTAGACAGATGAAGAGCCGGCTACATCGTAACCTCGGTCATG 99
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 3 IlelYsAlaGlnLeuLysGlnIleAspLysGlyLeuTyrArgGluLeuGlnProile 22
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 100 GATGAGCGCGGTTCCAGAGAGCAATATTGATCGCGAATCAAAACGGTCTGCTCA 159
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 23 GlnSerValGluLysGlnTyrIleTyrIleAsnAspGlnSerTyrIleAsnPheThSer 42
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 160 AACAAATTATTAGGCTCGCA-----AGCATAGACGCTTTGATCGATCGAGCC 207
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 43 AsnAspTyrLeuGlyIleGlyGlnValGluTyrGlnProGlnAsnPheLeuAspPheile 62
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 208 CAACAGCATTCAGCAATTTGGACAGGACGGTTCACGTTTAAACAGAGCAAT 267
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 63 LysThr-----TyrSerIleHisLeuSerSerSerArgLeuValSerGlyAsn 78
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 268 TCGGTCTGGCATGAAAGCTAGAAAGAAAGATTGCCAGCTTTAAACTGACAGAGCGGCC 327
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 79 SerValValTyrGlnGlnLeuGluGlnGluIleSerGluHisPheAsnPheGluAspAla 98
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 328 CTGCTGTTTCGAGCGGTTACTTGGCAATGTCGGTGTCTTCCTTCATCTTCCCGCAAGAAAG 387
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 99 LeuilePheAsnSerGlyTyrAspAlaAsnLeuAlaValPheAsnIlePheLysAsnAsn 118
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 388 GAAGATGCTCATTTTAAGTCAGCACCTCAATCATCAAGTATGATCGAGCGGTCGCGACTT 447
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 119 AsnIleValIlePheSerAspGlnGlnAsnHisAlaSerIleAspGlyIleLysLeu 138
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 448 TCTAAGGCTGCATACAGTGTATTTCGCGCATATTGATATGATCTTGAACCAAGCTG 507
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 139 SerGlyLeuSerLysValIleTyrGlnHisLeuAsnTyrAspAspLeuGluSerHisLeu 158
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 508 ---AATGAAACACAGCGGTTATACGCGCGGTTTTATCGTAACAGCGGAGTATTACAGATG 564
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 159 AlaArgHisThrAsnProAspValGlnLysValIleValSerAspSerValPheSerThr 178
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 565 GATGCGCAATCGCCCTCTTGATGATCATCTTCTACCTTGGGAAACGCTATCGCTTC 624
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 179 AsnGlyThrLysAlaAspIleAsnArgLeuValHisLeuLysGlnArgTyrAsnAlaile 198
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 625 GTGTCGTGTGATGATGCCACGCAACAGAGGTTTTGGCGGATTCGGGCAACAGGAGT 684
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 199 LeuilelleAspAlaSerHisSerLeuGlyLeu-----AsnLeuPhe 212
Db : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 685 GAATACTTGTGTTGTCCGACATTTGTTATCGGCACCTTAAGCAAAAGCTGTGGCGCG 744
```

```
Db 213 GluTyrHisAlaAspIle---AspIleValThrSerSerLysAlaTrpGlyAla 231
Qy 745 GAAGGAGGTTTTCGGGAGGATCAGCGTCTTCATCGACTTTTCTGCTGAACCATGCCAGA 804
Db 232 HisGlyGlyValIlePheSerSerLysAspLysAspLeuIleAsnLysGlyArg 251
Qy 805 ACATTATCTTCAAAACCGCTATTTCGCCAGCCAGCTGTGCGGTGCTCAGAGGCTTTC 864
Db 252 SerLeuileTyrSerSerSerLeuPro-----SerTyrHisLeuTyrPhe 266
Qy 865 -----ACATCATTTGAAGCCAGC-----AGGAAAAACGACAGCTT 900
Db 267 IleGlnValSerLeuGlnHisValIleGluAspThrTyrArgArgGluLysLeuAsnVal 286
Qy 901 TTATTTTCTTATATACG-----ATGATCAGACAGCTCTGAAGAT 942
Db 287 LeuSerGluTyrPheAsnHisGlnPheMetGluLeuPheProAspGlnProLeuSerAsn 306
Qy 943 ATGGGTTTATGTGTGAAGGAGATCACACACCGGATTTCTGTAGTCATTGGCGGATGCC 1002
Db 307 -----ThrProileLysAsnIleValCysAspSerLeu 317
Qy 1003 CATAAACGGTCTTATTGCTGAAAACTGACGGCAAGGAATTTATGCTCTGCCATT 1062
Db 318 AlaSerAlaGlnAlaGlnTyrAspMetLeuPheGluHisGlyIlePheValSerTyrLeu 337
Qy 1063 CGGCGCCCAACCGTTGCGCGCGGTGAAGCGGATT 1098
Db 338 ArgTyrProThrValSerGlnLeuThrLeuArgile 349

RESULT 7
US-11-055-822-644
; Sequence 644, Application US/11055822
; Publication No. US20050260707A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberkauer, Gregor
; TITLE OF INVENTION: CORNEBACTERIUM GLUTAMICUM GENES ENCODING
; FILE REFERENCE: BGI-121CPCN
; CURRENT APPLICATION NUMBER: US/11/055,822
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: 09/606,740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141,031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142,101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148,613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187,970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19930476.9
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931415.2
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931418.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1158
; SEQ ID NO 644
; LENGTH: 334
; TYPE: PRF
; ORGANISM: Corynebacterium glutamicum
US-11-055-822-644
```



```
Db 295 AspProGlnProGluSerValProIleAsnArgLeuValLysValGluGlyThrProLeu 314
Qy 2841 GAAGCGCTCAACGAAATTAACCGCTGATTTGTTTAAAGTGCTGGCGCTGTCCTGTTT 2900
Db 315 AlaAspAlaGluAspLeuAspTyrThrGluPheValArgThrValSerValAlaArgIle 334
Qy 2901 ATCAATCCATCAAAAGAAATTCGATTTCCGGAGAGAGAGAGAGTCAATCTCCGACATTTG 2960
Db 335 ThrMetProGlnSerTyrValArgLeuSerAlaGlyArgSerAsnMetProGluAlaMet 354
Qy 2961 CAGCATTAGGCTTTAGCCG---GCAAACTCCATTTTTCGAGACTACTTAACACT 3017
Db 355 GlnAlaMetCysPheMetAlaGlyAlaAsnSerIlePheTyrGlyAspLysLeuLeuThr 374
Qy 3018 GCCGGCAAGAGAGACGCGGATCATAAATGCTGAGTGAT----- 3059
Db 375 ThrGlyAsnProAspGluAspGlyAspArgIleLeuMetGluLysLeuAsnLeuTyrPro 394
Qy 3060 TTAGGCTTTGAAGTTGAATCA-----GTCGAAGAATGAAGGCTAGT 3101
Db 395 LeuGlnPheGluProGluGluValAlaGluValGluLysAlaSer 410

RESULT 9
US-10-657-4318
; Sequence 4318, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 4318
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-657-4318

Alignment Scores:
Pred. No.: 0.0155 Length: 339
Score: 115.00 Matches: 64
Percent Similarity: 39.1% Conservative: 56
Best Local Similarity: 20.8% Mismatches: 103
Query Match: 2.1% Indels: 84
DB: 6 Gaps: 17

US-10-681-086-1 (1-3156) x US-10-467-657-4318 (1-339)
Qy 265 AATTCGGTCTGGCAT-----GAAAGCTAGAAAG 294
Db 2 HisLysLeuThrHisIleSerAsnIleTyrThrThrArgProAlaGlnGluLeuAlaGln 21
Qy 295 AAGATTGCCACTTTAAACTGACAGAGCGCCCTGCTGTTTCGAGCGGTTACTTGCC 354
Db 22 LysLeuValLysHisSerPheAlaAspLysValPhePheCysAsnSerGlyAlaGluAla 41
Qy 355 AATTCGGTCTGCTTTCATCTCTGCCAGAA-----AAG 387
Db 42 AsnGluAlaAlaLeuLysLeuAlaArgLysTyrAlaArgAspArgPheGlyGlyLys 61
Qy 388 GAAGATGTCATTTTAAAGTGACCGCTCAATCATGCAAGTATGATC----- 432
Db 62 SerGluIleValAlaCysIleAsnSerPheHisGlyArgThrLeuPheThrValSerVal 81
```

```
Qy 433 -----GACGGTGGCGACTTCTTAAGGCTGATACAGTTGTT 468
Db 82 GlyGlyGlnProLysTyrSerLysAspTyrAlaProLeuProGlnGlyIleThr----- 99
Qy 469 TATCGGCATATTGATATGAATGAT-----CTTCAAAAACAGCTGAATGAACACAG 519
Db 100 -----HisValProPheAsnAspIleAlaAlaLeuGluAlaValGlyGlu----- 115
Qy 520 CGTTATCAGCGCCGTTTATCGTAAACAGAGGAGTATTACAGCATGGATGCACAAATGCC 579
Db 116 -----GlnThrCysAlaValIleLeuProIleGlnGlyGluSerGlyIleLeuPro 133
Qy 580 CCTCTTCATCAGATCATCTCCTGCGAAA-----CGTATCATGCTTCGTCG 627
Db 134 AlaThrAlaGluTyrLeuGlnAlaAlaArgLeuCysAspArgHisAsnAlaLeuLeu 153
Qy 628 GTCTTCATGATGCCACGCAACAGGAGTTTGGGGGATTCGGGACAA---GGAACGAGT 684
Db 154 IleLeuAspGluValGln---ThrGly---MetGlyHisThrGlyArgLeuPheAlaTyr 171
Qy 685 GAATACTTGGTGTTCCTCCGACATTTGTTATCGGCACCTTAAAGCAAGCTGTTGGCGCG 744
Db 172 GluHisTyrGlyValValProAspIleLeu-----SerSerAlaLysAlaLeuGlyCys 189
Qy 745 GAAGGAGGTTTGGCGGAGGATCAGCGCTCTTCATCGACTTTTGTCTG----- 792
Db 190 -----GlyPheProIleGlyThrMetLeuAlaThrGluLysIleAlaAlaPheGln 207
Qy 793 -----AACCATGCCAACAATTTATCTTCAAAACCGCTATTCCGCCAGCCAGCTGTGCG 846
Db 208 ProGlyThrHisGlySerThrPhe-----GlyGlyAsnProMetAlaCysAla 223
Qy 847 GCTGCTCAGCAGGCTTTCAACATCATTTGAAGCCAGCAGGAGGAAAAACGACAGCTTTTAT 906
Db 224 ValGlySerArgAlaPheAspIleIleAsnAlaPro---GluThrLeuHisAsnValArg 242
Qy 907 TCTTATATCAGCATGATCAGAACCATCTCTAGAAATATGGTTATGTGGTGAAGAGAT 966
Db 243 SerGlnGlyGlnLysLeuGlnThrAlaLeuAspLeuGly----- 256
Qy 967 CACACACGATTTATTCCTGTAGTCATTGGCGATGCCCATAAACGGTCTTATTTGCTGAA 1026
Db 257 -----ArgLysThrGlyLeuPheSerGln 264

Qy 1027 AAATCGCAGGCAAGGGAATT 1047
Db 265 ---ValArgGlyMetGlyLeu 270

RESULT 10
US-10-467-657-1860
; Sequence 1860, Application US/10467657
; Publication No. US20050260581A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: FONTANA Maria Rita
; APPLICANT: PIZZA Mariagrazia
; APPLICANT: MASNANI Vega
; APPLICANT: MONACI Elisabetta
; TITLE OF INVENTION: GONOCOCCAL PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/467,657
; CURRENT FILING DATE: 2003-08-11
; PRIOR APPLICATION NUMBER: GB-0103424.8
; PRIOR FILING DATE: 2001-02-12
; NUMBER OF SEQ ID NOS: 9218
; SOFTWARE: SeqWin99, version 1.04
; SEQ ID NO 1860
; LENGTH: 401
; TYPE: PRT
; ORGANISM: Neisseria gonorrhoeae
US-10-467-657-1860

Alignment Scores:
```









```
Db 255 -----AlaPheIleGlyArgCysArgValValProLeuArg 266
QY 940 AATATGGGT 948
Db 267 AsnMetGly 269
RESULT 14
US-10-995-561-776
; Sequence 776, Application US/10995561
; Publication No. US20050272054A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 776
; LENGTH: 5935
; TYPE: PR1
; ORGANISM: Homo sapiens
US-10-995-561-776
Alignment Scores:
Pred. No.: 0.556 Length: 5935
Score: 103.50 Matches: 95
Percent Similarity: 36.8% Conservative: 76
Best Local Similarity: 20.4% Mismatches: 167
Query Match: 1.9% Indels: 127
DB: 6 Gaps: 21
US-10-681-086-1 (1-3156) x US-10-995-561-776 (1-5935)
QY 1912 TCAGGTGGACTACTAGAGAAGTGTATCAAACTGCTAAATCGGTGAGAACCCCAACGTTTC 1971
Db 1783 AsnGlyGlyValAsnProGluProPheArgAlaThrGlnAsnVal-----Phe 1798
QY 1972 CACGATCGATTGTCGCTTATCGTAGAGAGCTGTTGTAATGAGGCAATGCGGT 2031
Db 1799 ThrArgGlnLeuCysLeuGluHisAspGluLysLeuValSerTyrLeuSerLeuLeuArg 1818
QY 2032 AATATTGCTGTCGCAAGGACCGACACATCTTCTACCATCTTGTACTGTACAGGTAGCAATG 2091
Db 1819 AsnIleGluMetArg-----ThrLysGlnIleGlnPro 1829
QY 2092 GCAGGTGCCATGTTGATTGTTGCTGCATCATCGCATCTGT----- 2130
Db 1830 LeuGluLeuAsnLeuAlaGluLeuGlnAspLeuLeuCysGlnAlaLysValLeuGluArg 1849
QY 2131 -----TATACGACGAGCGCTTCGCTTAACTGAAGCAGTTAAGCAATCAGAT 2178
Db 1850 GluLeuLysAspLeuThrThrLeuValSerGlnGluLeuCysValAsnGlnIle 1869
QY 2179 CTTCCTTCAGTTATGACCATCTGTCGAGTTCGTA-----ATGTCGTGTCACCTTTC 2232
Db 1870 IleSerGlnProGlnGluValProAlaGlnLeuLeuLysAlaLeuGluLysAspAlaLys 1889
QY 2233 GACTCTGAGAACTTCTGGATCTGCTAGAGAACTTCTGGATGGGATTCAGGAGTGGACA 2292
Db 1890 AsnLeuGlnLysSerLeuSerValSerAspThrTrpAsn----- 1903
QY 2293 GAACGACACGGATATATAGTGAGTGTCTCAAAACGCATACCATTTGAAATTCGAAAGCGC 2352
Db 1904 -----SerArgLeuLeu-HisPheGln-AsnAlaValG 1914
QY 2353 CGATTGACTTACCGGATGTTGAATGAGGAAACGCTGCTTGAAGCGCGGAGCGGCGC 2412
Db 1914 IuIleGluLysThrLysValLeuAsnGlnHisThrGlnLeuGluGly-----ArgLeuG 1932
QY 2413 ACGATCTGAAT-----ATCGGCACATATTGTATCGTGGCAAGCGGACAGGTCGCTCTA 2466
```

```
Db 1932 InAspLeuArgAlaTrpValGlyAsnLysAsnLysLeuAsnSerLysGly---SerA 1951
QY 2467 ACAGAGAAGTGGATCAGGTCTAGATGCGGTT-----CAGGAAA 2505
Db 1951 enSerGluIleAsp-----ValAspSerLeuAsnLeuCysLeuGlnGlnTyrGluAspL 1969
QY 2506 TTAAGACGACGATGACGACTGAAGATTGTGCATGCTCTTGACTG----- 2549
Db 1969 euLysGlnProMetAlaGluArgLysAlaGlnLeuAspAlaLeuAlaPheAspIleGlnP 1989
QY 2550 -----TTGAAGCCAGACGAGCGGCGAAGCGCTCAAGATG 2583
Db 1989 hePheIleSerGluHisAlaGlnAspLeuSerProGlnGlnAsnArgGlnMetLeu--- 2007
QY 2584 CAGGAGTAGACCGCTATATCATATAATTGTAATACGTCACAGAGAAACATTCCTCAACATCA 2643
Db 2008 -----ArgLeuLeuAsnGlnLeuGlnArgSerPheGlnAspIleL 2021
QY 2644 CAACCTCA----- 2651
Db 2021 euGluGlnThrAlaAlaGlnValAspAlaLeuGlnGlyHisLeuGlnGlnMetGluGlnG 2041
QY 2652 -----CATACATCGATGACAGACTCAATACGTTGAAATCGCAAAAGAA--- 2696
Db 2041 luAlaLeuValLysThrLeuGlnLysGlnGlnAsnThrCysHisGlnGlnLeuGluAspL 2061
QY 2697 -----TCGGGGCTGCTCCGCTGTTAGGCGCCCATTTATCGGGATG-----AAGGAGACGA 2745
Db 2061 euCysSerTrpValGlyGlnAlaGluArgAlaLeuAlaGlyHisGlnGlyArgThrG 2081
QY 2746 AACAGAGTGTCAATTGACATCGCCAAAGC-----TTGAAGGCTCTT-----GACG 2790
Db 2081 lnglnAspLeuSerAlaLeuGlnLysAsnGlnSerAspLeuLysAspLeuGlnAspAspI 2101
QY 2791 CGGATTCCATTCTCTGCAATTT-----TTGCATGCATTTGATGGCAGCGCTTAG 2841
Db 2101 leGlnAsnArgAlaThrSerPheAlaThrValLysAspIleGluGlyPheMetGluG 2121
QY 2842 AAGCGCTCAACGAATTAACCCGCTGTATTGTTTAAAGTGTGGCGCTGTCGTTTGA 2901
Db 2121 luAsnGlnThrLysLeuSerPro-----ArgGluLeuThrAlaLeuArgGluLysL 2138
QY 2902 TCAATCCATCAAAAGAAATTCGCATTTTCGAGAGAGAGAGGTCATTCCTCGCACATGCG 2961
Db 2138 euHisGlnAlaLysGlu-----GlnTyrGluAlaLeuG 2149
QY 2962 AGCCATTAGGCTTTAGCGCGCAAACTCCATTTTTCGCGAGACTACTTAACACTCGCG 3021
Db 2149 lnglnGluThrArgValAlaGlnLysGluLeu-----GluGluAlaValThrSerAlaL 2167
QY 3022 GGCAGAGGAGACGAGGATCATAAATGCTGAGTGATTTAGGCTTTGAAAGTTCAATCAG 3081
Db 2167 euGlnGlnGluThrGluLysSerLysAlaLysGluLeuAlaGluAsnLysLysI 2187
QY 3082 TCGAAGAAATG 3092
Db 2187 leAspAlaLeu 2190
RESULT 15
US-11-055-822-662
; Sequence 662, Application US/11055822
; Publication No. US20050260707A1
; GENERAL INFORMATION:
; APPLICANT: Pompejus, Markus
; APPLICANT: Kroger, Burkhard
; APPLICANT: Schroder, Hartwig
; APPLICANT: Zelder, Oskar
; APPLICANT: Haberhauser, Gregor
; TITLE OF INVENTION: CORYNEBACTERIUM GLUTAMICUM GENES ENCODING
; TITLE OF INVENTION: METABOLIC PATHWAY PROTEINS
; FILE REFERENCE: BGI-121CPCN
; CURRENT APPLICATION NUMBER: US/11/055,822
```

```
; CURRENT FILING DATE: 2005-02-11
; PRIOR APPLICATION NUMBER: 09/606,740
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/141,031
; PRIOR FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: 60/142,101
; PRIOR FILING DATE: 1999-07-02
; PRIOR APPLICATION NUMBER: 60/148,613
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/187,970
; PRIOR FILING DATE: 2000-03-09
; PRIOR APPLICATION NUMBER: DE 19930476.9
; PRIOR FILING DATE: 1999-07-01
; PRIOR APPLICATION NUMBER: DE 19931415.2
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931418.7
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931419.5
; PRIOR FILING DATE: 1999-07-08
; PRIOR APPLICATION NUMBER: DE 19931420.9
; PRIOR FILING DATE: 1999-07-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1158
; SEQ ID NO 662
; LENGTH: 348
; TYPE: PRT
; ORGANISM: Corynebacterium glutamicum
US-11-055-822-662

Alignment Scores:
Pred. No.: 0.239 Length: 348
Score: 102.00 Matches: 52
Percent Similarity: 26.4% Conservative: 34
Best Local Similarity: 22.0% Mismatches: 77
Query Match: 1.8% Indels: 73
DB: 7 Gaps: 11

US-10-681-086-1 (1-3156) x US-11-055-822-662 (1-348)
QY 2195 ACCATCTGTGGCAGTTCGTATGTCTGTCACTTCGACTCTGAGAACTTCGGAAAT 2254
Db 53 ThrValCysGlnGluAlaGlyCys-----ProAsnIleHisGluCysTrpGlu 68
QY 2255 CGCTAGAGAAATTCGGAAATCGGATTACAGAGTGGACAGACGCGATATATAGTGG 2314
Db 69 -SerArgGluAlaThrPheLeuIleGlyGlyAlaAsnCysSerArgarg----- 84
QY 2315 ATGTGTCAAAACGCATACCATTTTGAATTCGAAAGCGCGC----- 2354
Db 85 -Cys-----AapPheCysMetIleAsnSerAlaArgProGluProLeuAspArgGlyG1 102
QY 2355 -----ATTGAGTCTTACCG-----ATGCTGAATAAGGAAACGCTGCTT-- 2393
Db 102 uProLeuArgValAlaGluSerValArgGluMetGlnLeuAsnTyrSerThrIleThrG1 122
QY 2394 -----GAAGCGCGC-----AAGCG 2407
Db 122 yValThrArgAspAspLeuAspAspGluGluGlyAlaTrpLeuTyrSerGluValValArgLy 142
QY 2408 GCGGCACGATCTGAAT-----ATCGGCACATATTGTATCGTGGCAAGCGCAGAGTCC 2461
Db 142 sIleHisGluLeuAsnProHisThrGlyValGluAsnLeuValProAspPheSerGlyLy 162
QY 2462 GTCTAACAGAGAAGTGGATCGGTAGATCGGTTCAGGAATTAAGAGAGATATGG 2521
Db 162 sLys-----AspLeuLeuGlnGluValPheGluSer----- 172
QY 2522 ACTGAAGATTGTGCATGCTTTGGACTGTTGAAGCCAGACGACGAGCGGCTCAAAGA 2581
Db 173 -----ArgProGluVal----- 176
QY 2582 TGCAGGAGTAGACCGCTATAATCATTAATTTGAATACGTACACAGAAACCATTCAAACAT 2641
Db 2582 TGCAGGAGTAGACCGCTATAATCATTAATTTGAATACGTACACAGAAACCATTCAAACAT 2641
```

```
Db 177 -----PheAlaHisAsnValGluThrValProArgIlePheLysArgI1 191
QY 2642 CACAACCTCACATACATACGATGACAGATCAATACGGTTCAAAATCGCAAAAGATCGGG 2701
Db 191 eArgProAlaPheArgTyrGluArgSerLeuAspValIleArgGlnAlaArgAspPheG1 211
QY 2702 GCTGTCTCCGTGTTTCAGGCGCCATTATCGGATCAAGAGACAGAAACAGAGGATGTCATTGA 2761
Db 211 yLeuValThrLysSerAsnLeuIleLeuGlyMetGlyGluThrLysGluGluIleThrG1 231
QY 2762 CATGCCCAAAAGCTTGAAGGCTTTGACGGGATTCATTCCATTCCTGTG 2807
Db 231 uAlaLeuGlnAspLeuHisAspAlaGlyCysAspIleIleThrIle 246
```

Search completed: February 10, 2006, 04:07:17  
Job time : 57 secs

z1: /cyll2\_8/prodata/1/paa/0009/\_comb.pdf:

```

40      851 15.4 338 19 US-08-935-263-8      Sequence 8, Appli
41      843 15.3 392 1 PCT-US03-09585-26      Sequence 26, Appl
42      843 15.3 392 27 US-09-791-537-44892      Sequence 44892, A
43      843 15.3 392 33 US-10-369-493-23135      Sequence 23135, A
44      843 15.3 392 35 US-10-507-720-26      Sequence 26, Appl
45      843 15.3 392 47 US-60-360-039-23135      Sequence 23135, A

                                ALIGNMENTS

RESULT 1
US-09-791-537-52950
; Sequence 52950, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 52950
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-09-791-537-52950

Alignment Scores:
Pred. No.:      7,45e-168      Length:      389
Score:      1836.00      Matches:      359
Percent Similarity:      100.0%      Conservative:      1
Best Local Similarity:      99.7%      Mismatches:      0
Query Match:      33.3%      Indels:      0
DB:      27      Gaps:      0

US-10-681-086-1 (1-3156) x US-09-791-537-52950 (1-389)
QY      22 TTGAAGATTGATCTCTGTTAAACGAGCGGTTAGACAGATGAAAGAACGGCGGTACAT 81
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      1 MetLysIleAspSerTrpLeuAsnGluArgLeuAspArgMetLysGluAlaGlyValHis 20
QY      82 CTAACCTCGGTCATGATGAGCGCGCGTTCAGAGAGCAATATTGATGCGCAAAAT 141
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      21 ArgAsnLeuArgSerMetAspGlyAlaProValProGluArgAsnIleAspGlyGluAsn 40
QY      142 CAAACGGTCTGCTCTCAACCAATATTATTAGGCTCGCAAGCATAGACGTTTGATCGAT 201
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      41 GlnThrValTrpSerSerAsnAsnTyrLeuGlyLeuAlaSerAspArgLeuIleAsp 60
QY      202 GCAGCCCCAACAGATTGCAGCAATTGGGACAGGACGCGGTTACGTTTAAACGACA 261
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      61 AlaAlaGlnThrAlaLeuGlnpGlyThrGlySerSerGlySerArgLeuThrThr 80
QY      262 GCAATTCGGTTCGCATGAAAGAGTCAGAAAGAGATTCGCGCTTAAACTGCAGAA 321
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      81 GlyAsnSerValTrpHisGluLysLeuGluLysIleAlaSerPheLysLeuThrGlu 100
QY      322 GGGGCCCTGCTGTTTTCAGCGGTTACTTGGCCAATGTGCGGTGCTCTTTTCATCTTGCCA 381
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      101 AlaAlaLeuLeuPheSerSerGlyTyrLeuAlaAsnValGlyValLeuSerSerLeuPro 120
QY      382 GAAAGGAGATGTCTATTTAAGTGACCGACTCAATCATGCAAGTATGATCGACGGCTGC 441
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      121 GluLysGluAspValIleLeuSerAspGlnLeuAsnHisAlaSerMetIleAspGlyCys 140
QY      442 CGACTTCTTAGGCTGATACAGTCTTTATCGGCATATTGATGAATGATCTTCAAAAC 501
      ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db      141 ArgLeuSerLysAlaAspThrValValTyrArgHisIleAspMetAsnAspLeuGluAsn 160
QY      502 AAGCTGAATGAAACACAGCGTTATCAGCGCCGCTTTTATCGTAAACAGCGGAGTATTCAGC 561
```

```

      161 LysLeuAsnGluThrGlnArgTyrGlnArgArgPheIleValThrAspGlyValPheSer 180
      562 ATGATGAGCAACAATCGCCCTCTTGATCAGATCATCTCCTTGGAAACGTCATCATGCC 621
      181 MetAspGlyThrIleAlaProLeuAspGlnIleIleSerLeuAlaLysArgTyrHisAla 200
      622 TTCGTGTCGTTGATGATGCCACGACGAGGATTTTGGCGGATTCGGGACAAGGAACG 681
      201 PheValValValAspAspAlaHisAlaThrGlyValLeuGlyAspSerGlyGlnGlyThr 220
      682 AGTGAATACTTTGCTGTTTGTCCGACATTTTATCGGCACCTTTAAGCAAAAGCTGTTGCG 741
      221 SerGluTyrPheGlyValCysProAspIleValIleGlyThrLeuSerLysAlaValGly 240
      742 GCGAAGAGAGTTTGGCGAGGATCAGCGGTCTTCATCGACTTTTGTCTGAACCATGCC 801
      241 AlaGluGlyGlyPheAlaAlaGlySerAlaValPheIleAspPheLeuLeuAsnHisAla 260
      802 AGAACATTATCTTTCAACCGCTATTCCGCGCAGCAGCTGTGCGGCTGCACGAGGCT 861
      261 ArgThrPheIlePheGlnThrAlaIleProProlaserCysAlaAlaHisGluAla 280
      862 TTCAACATCATTTGAAGCCAGCAGCGGAAAAACGACAGCTTTTATTTTCTTATATCAGCATG 921
      281 PheAsnIleIleGluAlaSerArgGluLysArgGlnLeuLeuPheSerTyrIleSerMet 300
      922 ATCAGAACCATCTGAAGAATATCGGTATGTGTGTGTAAGGAGATCACACACCGATTATT 981
      301 IleArgThrSerLeuLysAsnMetGlyTyrValValLysGlyAspHisThrProIleIle 320
      982 CCTGTAGTCATTGGCGATGCCCATAAAAACGTCCTATTGCTGAAAAAAGCTGCAGGGCAAG 1041
      321 ProValValIleGlyAspAlaHisLysThrValLeuPheAlaGluLysLeuGlnGlyLys 340
      1042 GGAATTTATGCTCTGCCATTTCGCGCCGCAACCGTTGCGCGGTTGAAGCCGGATTGCA 1101
      341 GlyIleTyrAlaProAlaIleArgProThrValAlaProGlyGluSerArgIleArg 360

RESULT 2
US-10-369-493-23255
; Sequence 23255, Application US/10369493
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 23255
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Bacillus subtilis
US-10-369-493-23255

Alignment Scores:
Pred. No.:      7,45e-168      Length:      389
Score:      1836.00      Matches:      359
Percent Similarity:      100.0%      Conservative:      1
Best Local Similarity:      99.7%      Mismatches:      0
Query Match:      33.3%      Indels:      0
DB:      33      Gaps:      0

US-10-681-086-1 (1-3156) x US-10-369-493-23255 (1-389)
QY      22 TTGAAGATTGATCTCTGTTAAACGAGCGGTTAGACAGATGAAAGAACGGCGGTACAT 81
```

```
Db 1 MetLysileAspSerTrpLeuAsnGluArgLeuAspArgMetLysGluAlaGlyValHis 20
Qy 82 CGTAACCTCGCGTCAATGATGGAGCCGCGTTCAGAGAGGATATTTGATGGCGAAAT 141
Db 21 ArgAsnLeuArgSerMetAspGlyAlaProValProGluArgAsnIleAspGlyGluAsn 40
Qy 142 CAAACGGTCTCGTCTCTCAAAACAATTTATAGGCTCGCAAGCGATAGACGTTTCATCGAT 201
Db 41 GlnThrValTrpSerAsnAsnTyrLeuGlyLeuAlaSerAspArgLeuIleAsp 60
Qy 202 GCAGCCCAACAGCATTTGCAGCAATTTGGGACAGAGACGCGTTCACGTTTAACGACA 261
Db 61 AlaAlaGlnThrAlaLeuGlnGlnPheGlyThrGlySerSerGlySerArgLeuThr 80
Qy 262 GGCATTCGGTCTGCGCATGAAGCTAGAAAGAGATTCACGCTTTAACTCAGACAA 321
Db 81 GlyAsnSerValTrpHisGluLysLeuGluLysIleAlaSerPheLysLeuThrGlu 100
Qy 322 CGCGCCCTCGTCTTTTCGAGCGGTACTTTCGCGCAATGTCGCGTCTCTTCATCTTGCCA 381
Db 101 AlaAlaLeuLeuPheSerSerGlyTyrLeuAlaAsnValGlyValLeuSerSerLeuPro 120
Qy 382 GAAAGGAAGATGTCATTTTAAGTGACGAGTCAATCATGCAAGTATGATCGACGGCTGC 441
Db 121 GluLysGluAspValIleLeuSerAspGlnLeuAsnHisAlaSerMetIleAspGlyCys 140
Qy 442 CGACTTTCTAAGCGCTGATACAGTTGTTTATCGGCATATTTGATATGATGATCTTGAAG 501
Db 141 ArgLeuSerLysAlaAspThrValTyrArgHisIleAspMetAsnAspLeuGluAsn 160
Qy 502 AAGCTGTAATGAACACAGCGTTATCAGCGCGTTCATCGAAGAGATTCGCTTAACTCAGAC 561
Db 161 LysLeuAsnGluThrGlnArgTyrGlnArgArgPheIleValThrAspGlyValPheSer 180
Qy 562 ATGGATGSCACAATCGCCCTCTTGCATCAGATCATCTCACTTTCGAAACGCTATCATGCC 621
Db 181 MetAspGlyThrIleAlaProLeuAspGlnIleSerLeuAlaLysArgTyrHisAla 200
Qy 622 TTCGTGTCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 681
Db 201 PheValValValAspAspAlaHisAlaThrGlyValLeuGlyAspSerGlyGlnGlyThr 220
Qy 682 AGTCAATCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 741
Db 221 SerGluTyrPheGlyValCysProAspIleValIleGlyThrLeuSerLysAlaValGly 240
Qy 742 CGGAAGAGAGGTTTTCGCGCAGGATCAGCGGTCTTCATCGACTTTTTCGCTGAACCATGCC 801
Db 241 AlaGluGlyGlyPheAlaAlaGlySerAlaValPheIleAspPheLeuLeuAsnHisAla 260
Qy 802 AGACATTTATCTTTCAACCGCTATTCGCGCAGGAGTTCGCGGTGCTCAGAGGCT 861
Db 261 ArgThrPheIlePheGlnThrAlaIleProProAlaSerCysAlaAlaAlaHisGluAla 280
Qy 862 TTCACATCATTTGAAGCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 921
Db 281 PheAsnIleIleGluAlaSerArgGluLysArgGlnLeuLeuPheSerTyrIleSerMet 300
Qy 922 ATCAGAACCATCTCTGAAGATATGCGTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 981
Db 301 IleArgThrSerLeuLysAsnMetGlyTyrValValLysGlyAspPheIleThrProIleIle 320
Qy 982 CCTGTAGTCATTCGCGATGCCATTAACCGTCTTATTCGTTGTTGTTGTTGTTGTTGTTGTTGTT 1041
Db 321 ProValValIleGlyAspAlaHisLysThrValLeuPheAlaGluLysLeuGlnGlyLys 340
Qy 1042 GGAATTTATGCTCTGCGCATTCGCGCCGCAACCGTTGCGCGGTGGAAGCGCGATTCGA 1101
Db 341 GlyIleTyrAlaProAlaIleArgProThrValAlaProGlyGluSerArgIleArg 360
```

RESULT 3

US-60-360-039-23255

```
; Sequence 23255, Application US/60360039
; GENERAL INFORMATION:
```

```
; APPLICANT: Cao, Yongwei
```

```
; APPLICANT: Chen, Xianfeng
```

```
; APPLICANT: Goldman, Barry S.
```

```
; APPLICANT: Hinkle, Gregory J.
```

```
; APPLICANT: Slater, Steven C.
```

```
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
```

```
; FILE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
```

```
; FILE REFERENCE: 38-10(52052)A
```

```
; CURRENT APPLICATION NUMBER: US/60/360,039
```

```
; CURRENT FILING DATE: 2002-02-21
```

```
; NUMBER OF SEQ ID NOS: 47374
```

```
; SEQ ID NO 23255
```

```
; LENGTH: 389
```

```
; TYPE: PRT
```

```
; ORGANISM: Bacillus subtilis
```

```
; US-60-360-039-23255
```

```
Alignment Scores:
Pred. No.: 7,45e-168 Length: 389
Score: 1836.00 Matches: 359
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.7% Mismatches: 0
Query Match: 33.3% Indels: 0
DB: 47 Gaps: 0
```

```
US-10-681-086-1 (1-3156) x US-60-360-039-23255 (1-389)
```

```
Qy 22 TTGAAGATTCATCTCTCGTTTAAACGACGCGTTAGACAGATGAAGAAGCGCGGTACAT 81
Db 1 MetLysileAspSerTrpLeuAsnGluArgLeuAspArgMetLysGluAlaGlyValHis 20
Qy 82 CGTAACCTCGCGTCAATGATGGAGCCGCGTTCAGAGAGGATATTTGATGGCGAAAT 141
Db 21 ArgAsnLeuArgSerMetAspGlyAlaProValProGluArgAsnIleAspGlyGluAsn 40
Qy 142 CAAACGGTCTGCGCATGAAGCTAGAAAGAGATTCGCGCTTCGCAAGCGATAGACGTTTCATCGAT 201
Db 41 GlnThrValTrpSerAsnAsnTyrLeuGlyLeuAlaSerAspArgLeuIleAsp 60
Qy 202 GCAGCCCAACAGCATTTGCAGCAATTTGGGACAGAGACGCGTTCACGTTTAACGACA 261
Db 61 AlaAlaGlnThrAlaLeuGlnGlnPheGlyThrGlySerSerGlySerArgLeuThr 80
Qy 262 GGCATTCGGTCTGCGCATGAAGCTAGAAAGAGATTCGCGCTTTAACTCAGACAA 321
Db 81 GlyAsnSerValTrpHisGluLysLeuGluLysIleAlaSerPheLysLeuThrGlu 100
Qy 322 CGCGCCCTCGTCTTTTCGAGCGGTACTTTCGCGCAATGTCGCGTCTCTTCATCTTGCCA 381
Db 101 AlaAlaLeuLeuPheSerSerGlyTyrLeuAlaAsnValGlyValLeuSerSerLeuPro 120
Qy 382 GAAAGGAAGATGTCATTTTAAGTGACGAGTCAATCATGCAAGTATGATCGACGGCTGC 441
Db 121 GluLysGluAspValIleLeuSerAspGlnLeuAsnHisAlaSerMetIleAspGlyCys 140
Qy 442 CGACTTTCTAAGCGCTGATACAGTTGTTTATCGGCATATTTGATATGATGATCTTGAAG 501
Db 141 ArgLeuSerLysAlaAspThrValTyrArgHisIleAspMetAsnAspLeuGluAsn 160
Qy 502 AAGCTGTAATGAACACAGCGTTATCAGCGCGTTCATCGAAGAGATTCGCTTAACTCAGAC 561
Db 161 LysLeuAsnGluThrGlnArgTyrGlnArgArgPheIleValThrAspGlyValPheSer 180
Qy 562 ATGGATGSCACAATCGCCCTCTTGCATCAGATCATCTCACTTTCGAAACGCTATCATGCC 621
Db 181 MetAspGlyThrIleAlaProLeuAspGlnIleSerLeuAlaLysArgTyrHisAla 200
Qy 622 TTCGTGTCGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 681
Db 201 PheValValValAspAspAlaHisAlaThrGlyValLeuGlyAspSerGlyGlnGlyThr 220
```

```
QY 682 AGTAATACTTTGGTGTGTTTCCGACATTTGTTATCGGCACCTTAAGCAAAAGCTGTTGGC 741
Db 221 SerGluTyrPheGlyValCysProAspIleValIleGlyThrLeuSerLysAlaValGly 240
QY 742 GCGGAAGAGAGTTTGGCGCAGGATCAGCGGTCTTCATCGACTTTTCTGCTGAACCATGCC 801
Db 241 AlaGluGlyGlyPheAlaAlaGlySerAlaValPheIleAspPheLeuLeuAenHisAla 260
QY 802 AGACATTTACTTTTCAACCGCTATTCCGCCAGCAGCTGTCGGCTGCTCAGAGGCT 861
Db 261 ArgThrPheIlePheGlnThrAlaIleProAlaSerCysAlaAlaAlaHisGluAla 280
QY 862 TTCACATCATTTGAAGCCAGCAGGCAAAACACACAGCTTTTATTTCTTATATCAGCATG 921
Db 281 PheAenIleIleGluAlaSerArgGluLysArgGlnLeuLeuPheSerTyrIleSerMet 300
QY 922 ATCAGAACCAAGCTTGAAGAATATGGGTATGTGTGTAAGAGATCACACACCGATTATT 981
Db 301 IleArgThrSerLeuLysAenMetGlyTyrValValLysGlyAspHisThrProIleIle 320
QY 982 CCTGTAGTCATTGGCGATGCCATAAACGGTCCCTATTGCTGAAAACTCGCGGCAAG 1041
Db 321 ProValIleIleGlyAspAlaHisLysThrValLeuPheAlaGluLysLeuGlnGlyLys 340
QY 1042 GGAATTTATGCTCTGCCATTCCGCCGCCAACCGTTGCGCGGGTGAAGCCGGAATTCGA 1101
Db 341 GlyIleTyrAlaProAlaIleArgProThrValAlaProGlyGluSerArgIleArg 360

RESULT 4
US-60-638-099-24877
; Sequence 24877, Application US/60638099
; GENERAL INFORMATION:
; APPLICANT: Edgerton, Michael D
; TITLE OF INVENTION: Transgenic Plants with Enhanced Agronomic Phenotypes
; FILE REFERENCE: 38-21(53720)
; CURRENT APPLICATION NUMBER: US/60/638,099
; CURRENT FILING DATE: 2004-12-21
; NUMBER OF SEQ ID NOS: 48056
; SEQ ID NO 24877
; LENGTH: 389
; TYPE: PRT
; ORGANISM: Bacillus subtilis subsp. subtilis str. 168
US-60-638-099-24877
```

```
Alignment Scores:
Pred. No.: 7,45e-168 Length: 389
Score: 1836.00 Matches: 359
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.7% Mismatches: 0
Query Match: 33.3% Indels: 0
DB: 50 Gaps: 0
```

US-10-681-086-1 (1-3156) x US-60-638-099-24877 (1-389)

```
QY 22 TTGAAGATTGATCTCTGGTTAAACGAGCGGTAGACAGAATGAAGAAGCCGCGTACAT 81
Db 1 MetLysIleAspSerTrpLeuAsnGluArgLeuAspArgMetLysGluAlaGlyValHis 20
QY 82 CTAACCTCGGTCAATGATGAGCGCGCGGTTCAGAGAGAAATATTGATGGCGAAAT 141
Db 21 ArgAenLeuArgSerMetAspGlyAlaProValProGluArgAsnIleAspGlyGluAen 40
QY 142 CAAACGGTCTGCTCTCAACAAATTTATAGGCTCGCAGCGATAGAGCTTTGATCGAT 201
Db 41 GlnThrValTrpSerSerAsnAsnTyrLeuGlyLeuAlaSerAspArgLeuIleAsp 60
QY 202 GCAGCCCAACAGCATTTGGCAGCAAGTTCAGCGGTTCACGTTTAAACGACA 261
Db 61 AlaAlaGlnThrAlaLeuGlnPheGlyThrGlySerSerGlySerArgLeuThrThr 80
QY 262 GCAATTCGGTTCGCATGAAAGCTAGAAAGAGATTGCCAGCTTTAAACTGCAGAA 321
Db 81 GlyAenSerValTrpHisGluLysLeuGluLysLysIleAlaSerPheLysLeuThrGlu 100
```

```
QY 322 GCGGCCCTGCTGTTTTCGAGCGGTACTTGCCCAATGTCGCGTGTCTTTCATCTTGCCA 381
Db 101 AlaAlaLeuLeuPheSerSerGlyTyrLeuAlaAenValGlyValLeuSerSerLeuPro 120
QY 382 GAAAGAAGATGTCATTTTAAAGTACAGCTCAATCATGCAAGTATGATCGAGCGGTGC 441
Db 121 GluLysGluAspValIleLeuSerAspGlnLeuAenHisAlaSerMetIleAspGlyCys 140
QY 442 CGACTTTCTAAGGCTGATACAGTTGTTTATCGGCATATTGATATGATGATCTTGAAC 501
Db 141 ArgLeuSerLysAlaAspThrValValTyrArgHisIleAspMetAenAspLeuGluAen 160
QY 502 AAGCTGAATGAAACACACAGCGTTATCAGCGCGTATTATCGTAACACAGCGAGTATTTCAGC 561
Db 161 LysLeuAenGlnThrGlnArgTyrGlnArgArgPheIleValThrAspGlyValPheSer 180
QY 562 ATGATGCGCAATTCGCCCTCTTTGATCAGATCATCTCATCTTGGCAACGCTATCATGCC 621
Db 181 MetAspGlyThrIleAlaProLeuAspGlnIleIleSerLeuAlaLysArgTyrHisAla 200
QY 622 TTCGTGCTGCTGATGATGCCACGACAGGAGTTCGCGGATTCGGGACAAAGGACG 681
Db 201 PheValValAspAspAlaHisAlaThrGlyValLeuGlyAspSerGlyGlnGlyThr 220
QY 682 AGTCAATACTTTGCTGTTTGTCCGACATTTGTTATCGGCACCTTAAAGCAAAAGCTGTTGGC 741
Db 221 SerGluTyrPheGlyValCysProAspIleValIleGlyThrLeuSerLysAlaValGly 240
QY 742 GCGGAAGAGGTTTGGCGCAGGATCAGCGGTCTTCATCGACTTTTCTGCTGAACCATGCC 801
Db 241 AlaGluGlyGlyPheAlaAlaGlySerAlaValPheIleAspPheLeuLeuAenHisAla 260
QY 802 AGACATTTACTTTTCAACCGCTATTCCGCCAGCAGCTGTCGGCTGCTCAGAGGCT 861
Db 261 ArgThrPheIlePheGlnThrAlaIleProAlaSerCysAlaAlaAlaHisGluAla 280
QY 862 TTCACATCATTTGAAGCCAGCAGGCAAAACACACAGCTTTTATTTCTTATATCAGCATG 921
Db 281 PheAenIleIleGluAlaSerArgGluLysArgGlnLeuLeuPheSerTyrIleSerMet 300
QY 922 ATCAGAACCAAGCTTGAAGAATATGGGTATGTGTGTAAGAGATCACACACCGATTATT 981
Db 301 IleArgThrSerLeuLysAenMetGlyTyrValValLysGlyAspHisThrProIleIle 320
QY 982 CCTGTAGTCATTGGCGATGCCATAAACCGTCTTATTTGCTGAAAACTCGCGGCAAG 1041
Db 321 ProValIleIleGlyAspAlaHisLysThrValLeuPheAlaGluLysLeuGlnGlyLys 340
QY 1042 GGAATTTATGCTCTGCCATTCCGCCGCCAACCGTTGCGCGGGTGAAGCCGGAATTCGA 1101
Db 341 GlyIleTyrAlaProAlaIleArgProThrValAlaProGlyGluSerArgIleArg 360
```

#### RESULT 5

US-60-638-099-43392  
; Sequence 43392, Application US/60638099  
; GENERAL INFORMATION:

; APPLICANT: Edgerton, Michael D  
; TITLE OF INVENTION: Transgenic Plants with Enhanced Agronomic Phenotypes  
; FILE REFERENCE: 38-21(53720)  
; CURRENT APPLICATION NUMBER: US/60/638,099  
; CURRENT FILING DATE: 2004-12-21  
; NUMBER OF SEQ ID NOS: 48056  
; SEQ ID NO 43392  
; LENGTH: 371  
; TYPE: PRT  
; ORGANISM: Bacillus subtilis  
US-60-638-099-43392

Alignment Scores:  
Pred. No.: 2.36e-152 Length: 371  
Score: 1676.00 Matches: 330  
Percent Similarity: 93.3% Conservative: 6



```
Best Local Similarity: 91.7% Mismatches: 6
Query Match: 30.4% Indels: 18
DB: 50 Gaps: 1

US-10-681-086-1 (1-3156) x US-60-638-099-43392 (1-371)

QY 22 TTGAAGATTGATTCCTGGTTAAACGACGGGTGACAGAAATGAAGAAGCCGCGTACAT 81
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MetLysileAspSerTrpLeuAenAspArgLeuAspIleAlaLysGluAlaGlyValHis 20

QY 82 CGTAACCTGGCGTCAATGGATGAGGCCCGTCCAGAGAGGAATATTGATGCGGAAT 141
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 21 ArgAsnLeuArgSerMetAsnGlyAlaProValProGluArgAsnIleAspGlyGluAsn 40

QY 142 CAAACGCTGCTGCTCTCAACAATTTATGGGCTCGCAAGCATAGACGCTTTGATCGAT 201
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 41 GlnThrValTrpSerAsnAsnTrpLeuGlyLeuAlaSerAspArgLeuIleAsp 60

QY 202 GCAGCCCAACAGCATTGCGCAATTTGGGACAGGACGCGGTTCACCTTTAAACGACA 261
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 AlaAlaGlnThrAlaLeuGlnPheGlyThrGlySerSerGlySerArgLeuThrThr 80

QY 262 GGCATTGCTGGCATGAAGCTAGAAAGAGATTCGCAGCTTTAACTGACAGAA 321
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 81 GlyAsnSerValTrpHisGluLysLeuGluLysIleAlaSerPheLysArgThrGlu 100

QY 322 CGGCGCTGCTGTTTCGAGCGGTACTTGCCCAATGTCGTGTCCTTCATCTTGCCCA 381
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 101 AlaAlaLeuLeuPheSerSerGlyTrpLeuAlaAsnValGlyValLeuSerSerLeuPro 120

QY 382 GAAAGGAAGATGTCATTTAAAGTACGAGTCAATCATGCAAGTATGATCGAGCGTGC 441
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 GluLysGluAspValIleLeuSerAspGlnLeuAsnHisAlaSerIleIleAspGlyCys 140

QY 442 CGACTTTCTAAGCTGATACAGTTGTTTATCGGCATATTGATGATGATGATCTTGAAC 501
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 141 ArgLeuSerLysAlaAspThrValValTrpArgHisIleAspMetAsnAspLeuGluAsn 160

QY 502 AAGCTGAATCAACACAGCGTTATCAGCGCGTTCATCGTAAAGAGGATGATTCAGC 561
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 161 LysLeuAsnGluThrGlnArgGlyArgGlyValThrAspGlyValPheSer 180

QY 562 ATGATCGGCAATCGCCCTCTTGATCAGATCATCTCACTTGCAGAACGCTATCATGCC 621
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 MetAspGlyThrIleValProLeuAspGlnIleSerLeuAlaLysArgTrpHisAla 200

QY 622 TTCGTGCTGTGATGATGCCACGCAACAGGAGTTTGGCGGATTCGGGACAGGAACG 681
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 201 PheValValAspAspAlaHisAlaThrGlyValLeuGlyAspSerGlyArgGlyThr 220

QY 682 AGTGAATACTTGTGTTGTCGCGCATTTGTTATCGCACCTTAAGCAAGCTGTGGC 741
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 221 SerGluTrpPheGlyValTrpProAspIleValIleGlyThrLeuSerLysAlaValGly 240

QY 742 GCGAAGAGGATTTGCGGAGGATCAGCGGTCTTCATCGACTTTTTCGCAACCATGCC 801
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 241 ThrGluGlyGlyPheAlaAlaGlySerAlaValPheIleAspPheLeuLeuAsnHisAla 260

QY 802 AGAACATTATCTTTTCAACCGCTATTCCGCGACGAGCTGTGCGGTCTGCTCAGAGCT 861
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 261 ArgThrPheIlePheGlnThrAlaIleProProAlaSerCysAlaAlaHisGluAla 280

QY 862 TTCACATCATTTGAAGCCAGCGGAAACGACAGCTTTTATTTCTTATATCAGCATG 921
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 281 PheAsn-----282

QY 922 ATCAGAACCACTCTCAAGAATATGGTGTATGTTGTTGAAGAGATCATCACACCGATTAT 981
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 283 IleArgThrSerLeuLysAsnMetGlyTrpValValLysGlyAspHisThrProIle 302

QY 982 CCTGTAGTCATTGCGCATGCCCAATAAACCGCTCTATTGCTGAAAACTGCGGCAAG 1041
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 303 ProValValIleGlyAspAlaHisLysThrValIlePheAlaGluLysLeuGlnGlyLys 322
```

## RESULT 6

```
US-08-827-356-3622
; Sequence 3622, Application US/08827356
; GENERAL INFORMATION:
; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Miller
; APPLICANT: Robert A. S. Hare
; APPLICANT: Karen J. Shaw
; TITLE OF INVENTION: STAPHYLOCOCCUS AUREUS RELATED
; NUMBER OF SEQUENCES: 5574
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Schering-Plough Corporation
; STREET: 2000 Galloping Hill Road
; CITY: Kenilworth
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07033-0530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/827,356
; FILING DATE: 01-APR-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/014,477
; FILING DATE: 01-APR-1996
; APPLICATION NUMBER: 60/016,743
; FILING DATE: 02-MAY-1996
; APPLICATION NUMBER: 60/020,016
; FILING DATE: 14-JUN-1996
; INFORMATION FOR SEQ ID NO: 3622:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 279 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: YES
; ORIGINAL SOURCE:
; ORGANISM: Staphylococcus aureus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1...279
US-08-827-356-3622

Alignment Scores:
Pred. No.: 9,22e-132 Length: 279
Score: 1463.00 Matches: 279
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 26.5% Indels: 0
DB: 18 Gaps: 0

US-10-681-086-1 (1-3156) x US-08-827-356-3622 (1-279)

QY 1501 CTGAATAAGAACGCGTCTCCAAATATCTTTATAGAAAAGCAATCTAAATATATCT 1560
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 LeuAsnLysAsnGlyAlaLeuGlnIlePheLeuPheArgLysAlaAsnLeuLysLeuSer 20

QY 1561 GAAAGGAATGAGAATAGTGAATGGAACCAATATATATAGTACAGAGAAAGAAATGAAG 1620
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 21 GluLysGlyMetArgIleValAsnGlyProIleIleMetThrArgGluArgMetLys 40

QY 1621 ATTGTTTCATGAATTAAGGAACGAATATTCGATTAATATTCGGGATGATGTTAAGCTATT 1680
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 41 IleValHisGluIleLysGluArgIleLeuAspLysTrpGlyAspAspValLysAlaIle 60
```

```
Qy 1661 GGTGTTTATGCTCTCTTGGTCGCAGACTGATGGGCCCTATTTCGGATATTGAGATGATG 1740
Db 61 GlyValTyrGlySerLeuGlyArgGlnThrAspGlyProTyrSerAspIleGluMetMet 80
Qy 1741 TGTGTCATGTCAACAGAGAAACGAGAGTTCAGCCATGAATGCAACACCGGTGAGTGAAG 1800
Db 81 CysValMetSerThrGluGluAlaGluPheSerHisGluTrpThrThrGlyGluTrpLys 100
Qy 1801 GTGAAGTGAATTTGATAGGAGAGAGATTCTCTAGATTATGCAATCTCAGGTGGAATCA 1860
Db 101 ValGluValAsnPheAspSerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSer 120
Qy 1861 GATTGGCCGCTTACACATGGTCAATTTTCTCTATTTCGCCGATTTATGATTCAGGTGGA 1920
Db 121 AsprTrpProLeuThrHisGlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGly 140
Qy 1921 TACTTAGAAGAAGTGTATCAAACTGCTAAATCGGTAGAGGCCCAACGTTCCACGATGCG 1980
Db 141 TyrLeuGluLysValTyrGlnThrAlaLysSerValGluAlaGlnThrPheHisAspAla 160
Qy 1981 ATTTGTGCCCTTATCGTAGAGAGCTGTTTGAATATGAGGCAAAATGGCGTAATTCGT 2040
Db 161 IleCysAlaLeuIleValGluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArg 180
Qy 2041 GTGCAAGGACGCAACAATTTCTACCATCTTGACTGTACAGGTAGCAATGCGAGGTGCC 2100
Db 181 ValGlnGlyProThrThrPheLeuProSerLeuThrValGlnValAlaMetAlaGlyAla 200
Qy 2101 ATGTTGATTGCTGTCATCATCGCATCTGTTTATACGACGAGCGCTTCGGTCTTAACTGAA 2160
Db 201 MetLeuIleGlyLeuHisHisArgIleCysTyrThrThrSerAlaSerValLeuThrGlu 220
Qy 2161 GCAGTTAAGCAATCAGATCTTCCTTCAGGTATGACCATCTGTGCCAGTTCGTAATGTCT 2220
Db 221 AlaValLysGlnSerAspLeuProSerGlyTyrAspHisLeuCysGlnPheValMetSer 240
Qy 2221 GGTCAACTTCCTGACTCTGAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATT 2280
Db 241 GlyGlnLeuSerAspSerGluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIle 260
Qy 2281 CAGGAGTGGACAGACGACGAGATATATAGTGGATGTGTCAAAACGCAATCCATTT 2337
Db 261 GlnGluTrpThrGluArgHisGlyTyrIleValAspValSerLysArgIleProPhe 279

RESULT 7
US-10-681-529-6706
; Sequence 6706, Application US/09611529
; GENERAL INFORMATION:
; APPLICANT: George H. Shimer, Jr.
; APPLICANT: George H. Miller
; APPLICANT: Roberta S. Hare
; APPLICANT: Karen J. Shaw
; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods
; FILE REFERENCE: 1034/1C963U51
; CURRENT APPLICATION NUMBER: US/09/611,529
; CURRENT FILING DATE: 2000-06-30
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: US 09/417,811
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/353,718
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/266,557
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,556
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,555
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,542
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,541
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/037,934
; PRIOR FILING DATE: 1998-03-10
```

```
; PRIOR APPLICATION NUMBER: US 09/036,720
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,338
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,334
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,221
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,137
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,082
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,081
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,079
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/035,913
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/035,744
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 08/827,356
; PRIOR FILING DATE: 1997-04-01
; PRIOR APPLICATION NUMBER: US 08/831,156
; PRIOR FILING DATE: 1997-04-01
; PRIOR APPLICATION NUMBER: US 60/014,477
; PRIOR FILING DATE: 1996-04-01
; PRIOR APPLICATION NUMBER: US 60/016,743
; PRIOR FILING DATE: 1996-05-02
; PRIOR APPLICATION NUMBER: US 60/020,016
; PRIOR FILING DATE: 1996-06-14
; NUMBER OF SEQ ID NOS: 7451
; SEQ ID NO 6706
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-611-529-6706

Alignment Scores:
Pred. No.: 9,22e-132 Length: 279
Score: 1463.00 Matches: 279
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 26.5% Indels: 0
DB: Gaps: 0

US-10-681-086-1 (1-3156) x US-09-611-529-6706 (1-279)
Qy 1501 CTGAATAAGAACGGTGTCTCCAAATATTTTATTTAGAAAAGCAAAATCTAAAAATTATCT 1560
Db 1 LeuAsnLysAsnGlyAlaLeuGlnIlePheLeuPheArgLysAlaAsnLeuLysLeuSer 20
Qy 1561 GAAAGGGATGAGATAGTGAATGAGCCAAATATATATAGTAGAGAGAAAGAAATGAAG 1620
Db 21 GluLysGlyMetArgIleValAsnGlyProIleIleMetThrArgGluGluArgMetLys 40
Qy 1621 ATTGTTTCATGAAATTAAGGAACGAATATTCGATAAATATGGGATGATGTTAAGGCTATT 1680
Db 41 IleValHisGluIleLysGluArgIleLeuAspLysTyrGlyAspAspValLysAlaIle 60
Qy 1681 GGTGTTTATGGCTCTCTTGGTCGTGACACTGATGGGCCCTATTTCGGATATTGAGATGATG 1740
Db 61 GlyValTyrGlySerLeuGlyArgGlnThrAspGlyProTyrSerAspIleGluMetMet 80
Qy 1741 TGTGTCATGTCAACAGAGAGAGAGTTCAGCCATGAATGCAACACCGGTGAGTGAAG 1800
Db 81 CysValMetSerThrGluGluAlaGluPheSerHisGluTrpThrThrGlyGluTrpLys 100
Qy 1801 GTGGAAGTGAATTTGATAGGAGAGAGATTCATAGATTATGCAATCTCAGGTGGAATCA 1860
Db 101 ValGluValAsnPheAspSerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSer 120
Qy 1861 GATTGGCCGCTTACACATGGTCAATTTTCTCTATTTCGCCGATTTATGATTCAGGTGGA 1920
```

Db 121 AspTrpProLeuThrHisGlyGlnPheSerIleLeuProIleTyrAspSerGlyGly 140  
QY 1921 TACTTAGAAGAGTGTATCAAACTGCTAAATCGGTAGAGCCCAAAAGCTTCCACGATGCG 1980  
Db 141 TyrLeuGluLysValTyrGlnThrAlaLysSerValGluAlaGlnThrPheHisAspAla 160  
QY 1981 ATTTGTGCGCTTATCGTAGAAGAGCTGTTTGAATATATGCGGCAAAATGGCGTAATATTCGT 2040  
Db 161 IleCysAlaLeuIleValGluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArg 180  
QY 2041 GTCAGAGGACCGACACATTTCTACCATCTTCCAGTGTACAGGTAGCAATGGCAGGTGCC 2100  
Db 181 ValGlnGlyProThrPheLeuProSerLeuThrValGlnValAlaMetAlaGlyAla 200  
QY 2101 ATGTTGATGCTGTCATCATCTGTCATCTGTTATACGACGAGCGCTTCGGTCTTAACTGAA 2160  
Db 201 MetLeuIleGlyLeuHisArgIleCysTyrThrSerAlaSerValLeuThrGlu 220  
QY 2161 GCAGTTAAGCAATCAGATCTTCTCAGGTTATGACCATCTGCGCAGTTCGTAATGTCT 2220  
Db 221 AlaValLysGlnSerAspLeuProSerGlyTyrAspHisLeuCysGlnPheValMetSer 240  
QY 2221 GGTCAACTTCCGACCTCGAAGACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGATT 2280  
Db 241 GlyGlnLeuSerAspSerGluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIle 260  
QY 2281 CAGGAGTGCAGACAGAACACACGCGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337  
Db 261 GlnGluTrpThrGluArgHisGlyTyrIleValAspValSerLysArgIleProPhe 279

## RESULT 8

US-09-950-084-6706  
; Sequence 6706, Application US/09950084  
; GENERAL INFORMATION:  
; APPLICANT: George H. Shimer, Jr.  
; APPLICANT: George H. Miller  
; APPLICANT: Roberta S. Hare  
; APPLICANT: Karen J. Shaw  
; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods  
; FILE REFERENCE: 1034/1C963US2  
; CURRENT APPLICATION NUMBER: US/09/950,084  
; CURRENT FILING DATE: 2001-09-10  
; PRIOR APPLICATION NUMBER: US 09/417,811  
; PRIOR FILING DATE: 1999-10-14  
; PRIOR APPLICATION NUMBER: US 09/353,718  
; PRIOR FILING DATE: 1999-07-14  
; PRIOR APPLICATION NUMBER: US 09/266,557  
; PRIOR FILING DATE: 1999-03-11  
; PRIOR APPLICATION NUMBER: US 09/266,556  
; PRIOR FILING DATE: 1999-03-11  
; PRIOR APPLICATION NUMBER: US 09/266,555  
; PRIOR FILING DATE: 1999-03-11  
; PRIOR APPLICATION NUMBER: US 09/266,542  
; PRIOR FILING DATE: 1999-03-11  
; PRIOR APPLICATION NUMBER: US 09/266,541  
; PRIOR FILING DATE: 1999-03-11  
; PRIOR APPLICATION NUMBER: US 09/037,934  
; PRIOR FILING DATE: 1998-03-10  
; PRIOR APPLICATION NUMBER: US 09/036,720  
; PRIOR FILING DATE: 1998-03-06  
; PRIOR APPLICATION NUMBER: US 09/036,338  
; PRIOR FILING DATE: 1998-03-06  
; Remaining prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 7451  
; SEQ ID NO 6706  
; LENGTH: 279  
; TYPE: PRT  
; ORGANISM: Staphylococcus aureus  
US-09-950-084-6706

Alignment Scores: 9.22e-132 Length: 279  
Pred. No.: 1463.00 Matches: 279  
Score:

Percent Similarity: 100.0% Conservative: 0  
Best Local Similarity: 100.0% Mismatches: 0  
Query Match: 26.5% Indels: 0  
DB: 29 Gaps: 0  
US-10-681-086-1 (1-3156) x US-09-950-084-6706 (1-279)  
QY 1501 CTGAATAAGACCGTCTCTCCAAATATTCTTATTTAGAAAACAAATCTTAAATATATCT 1560  
Db 1 LeuAsnLysAsnGlyAlaLeuGlnIlePheLeuPheArgLysAlaAsnLeuLysLeuSer 20  
QY 1561 GAAAGGGATCAGATAGTGAATGACCAATAATTAATGACTAGACAGAAAGAAAGTGAAG 1620  
Db 21 GluLysGlyMetArgIleValAsnGlyProIleIleMetThrArgGluGluArgMetLys 40  
QY 1621 ATTGTTTCATGAAATAAGGAAACGAATATTGGATAAATATGGGATGATGTTAAGGCTATT 1680  
Db 41 IleValHisGluIleLysGluArgIleLeuAspLysTyrGlyAspValLysAlaIle 60  
QY 1681 GGTGTTTATGGCTCTCTTGGTCTCAGACTGATGGCCCTATTTCGATATTGAGATGATG 1740  
Db 61 GlyValTyrGlySerLeuGlyArgGlnThrAspGlyProTyrSerAspIleGluMetMet 80  
QY 1741 TGTGTCATCTCAACAGAGGACGAGATTTCAGCCATGAATGGACACCCGCTGAGTGAAG 1800  
Db 81 CysValMetSerThrGluGluAlaGluPheSerHisGluTrpThrThrGlyGluTrpLys 100  
QY 1801 GTGGAAGTCAATTTTGTATAGCGAAGAGATTCTACTAGATTATGTCATCTCAGGTGGAATCA 1860  
Db 101 ValGluValAsnPheAspSerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSer 120  
QY 1861 GATTGCGCCCTTACACATGTCATATTTTCTTATTTTCCGATTTATGATTCAGGTGGA 1920  
Db 121 AspTrpProLeuThrHisGlyGlnPheSerIleLeuProIleTyrAspSerGlyGly 140  
QY 1921 TACTTAGAAGAGTGTATCAAACTGCTAAATCGGTAGAGCCCAAAAGCTTCCACGATGCG 1980  
Db 141 TyrLeuGluLysValTyrGlnThrAlaLysSerValGluAlaGlnThrPheHisAspAla 160  
QY 1981 ATTTGTGCGCTTATCGTAGAAGAGCTGTTTGAATATGACGCAAAATGGCGTAATATTCGT 2040  
Db 161 IleCysAlaLeuIleValGluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArg 180  
QY 2041 GTCAGAGGACCGACACATTTCTACCATCTTCCAGTGTACAGGTAGCAATGGCAGGTGCC 2100  
Db 181 ValGlnGlyProThrPheLeuProSerLeuThrValGlnValAlaMetAlaGlyAla 200  
QY 2101 ATGTTGATGCTGTCATCATCTGTCATCTGTTATACGACGAGCGCTTCGGTCTTAACTGAA 2160  
Db 201 MetLeuIleGlyLeuHisArgIleCysTyrThrSerAlaSerValLeuThrGlu 220  
QY 2161 GCAGTTAAGCAATCAGATCTTCTCAGGTTATGACCATCTGCGCAGTTCGTAATGTCT 2220  
Db 221 AlaValLysGlnSerAspLeuProSerGlyTyrAspHisLeuCysGlnPheValMetSer 240  
QY 2221 GGTCAACTTCCGACCTCTGAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGATT 2280  
Db 241 GlyGlnLeuSerAspSerGluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIle 260  
QY 2281 CAGGAGTGCAGACAGAACACGCGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337  
Db 261 GlnGluTrpThrGluArgHisGlyTyrIleValAspValSerLysArgIleProPhe 279  
RESULT 9  
US-10-914-020-6706  
; Sequence 6706, Application US/10914020  
; GENERAL INFORMATION:  
; APPLICANT: George H. Shimer, Jr. et al.  
; TITLE OF INVENTION: Staphylococcus aureus Related Compositions and Methods  
; FILE REFERENCE: 1034/1C963US1  
; CURRENT APPLICATION NUMBER: US/10/914,020  
; CURRENT FILING DATE: 2004-08-06  
; PRIOR APPLICATION NUMBER: US 09/417,811

```
; PRIOR FILING DATE: 1999-10-14
; PRIOR APPLICATION NUMBER: US 09/353,718
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/266,557
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,556
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,555
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,542
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/266,541
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 09/037,934
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: US 09/036,720
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 09/036,338
; PRIOR FILING DATE: 1998-03-06
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 7451
; SEQ ID NO 6706
; LENGTH: 279
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-10-914-020-6706

Alignment Scores:
Pred. No.: 9,22e-132 Length: 279
Score: 1463.00 Matches: 279
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 26.5% Indels: 0
DB: 39 Gaps: 0

US-10-681-086-1 (1-3156) x US-10-914-020-6706 (1-279)

QY 1501 CTGAATAGACGGTCTCCCAATATCTTATTAGAAAGCAATCTAATATTCT 1560
Db 1 LeuAnLysAeNGlyAlaLeuGlnIlePheLeuPheArgLysAlaAenLeuLysLeuSer 20

QY 1561 GAAAGGGAATCAGAAATAGTGAATGAGCAATATATGACTAGAGAAGAAAGAAATGAAG 1620
Db 21 GluLysGlyMetArgIleValAsnGlyProIleIleMetThrArgGluGluArgMetLys 40

QY 1621 ATTGTTGATGAATTAAGGAACGAATATGGAATAATATGGGATGATGTTAAGGCTATT 1680
Db 41 IleValHisGluIleLysGluArgIleLeuAspLysTyrGlyAspAspValLysAlaIle 60

QY 1681 GGCTGTTATGCTCTCTGGTCTGCTCAGACTGATGGCCCTATTTCGGATATTGAGATGATG 1740
Db 61 GlyValTyrGlySerLeuGlyArgGlnThrAspGlyProTyrSerAspIleGluMetMet 80

QY 1741 TGTGTCATGTCAACAGAGGAAGCAGAGTTAGCCATGAATGACAAACCGGTGAGTGAAG 1800
Db 81 CysValMetSerThrGluGluAlaGluPheSerHisGluIleThrThrGlyGluIleTrpLys 100

QY 1801 GTGAAGTGAATTTGATAGCAAGAGATTTCTACTAGATTATGCAATCTCAGGTGGAATCA 1860
Db 101 ValGluValAenPheAspSerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSer 120

QY 1861 GATTGGCCGCTACACATGGTCAATTTTCTCTATTTCGGATTTATGATTGAGTGA 1920
Db 121 AspTrpProLeuThrHisGlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGly 140

QY 1921 TACTTAGAAGTGTATCAAACTGCATAATCGGTAGAACCCAAACGTTCCAGATGCG 1980
Db 141 TyrLeuGluLysValTyrGlnThrAlaLysSerValGluAlaGlnThrPheHisAspAla 160

QY 1981 ATTTGTGCCCTTATCGTAGAAGAGCTGTTTGAATATATCGAGCAAAATCGCGTAATATTCTG 2040
Db 161 IleCysAlaLeuIleValGluGluLeuPheGluTyrAlaGlyLysTyrArgAsnIleArg 180
```

```
QY 2041 GTCAAGGACCGCAACACATTTTACCATTCTTACTGCTACAGGTAGCAATGCGAGTGCC 2100
Db 181 ValGlnGlyProThrThrPheLeuProSerLeuThrValGlnValAlaMetAlaGlyAla 200

QY 2101 ATGTTGATTGGTCTGCATCATCTGCATCTGTTATACGACGAGCGCTTCGGTCTTAACGTAA 2160
Db 201 MetLeuIleGlyLeuHisArgIleCysTyrThrThrSerAlaSerValLeuThrGlu 220

QY 2161 GCAGTTAAGCAATCAGATCTTCTTACAGTTATGACCATCTGTGCCAGTTCGTAATGCTCT 2220
Db 221 AlaValLysGlnSerAspLeuProSerGlyTyrAspHisLeuCysGlnPheValMetSer 240

QY 2221 GGTCAACTTTCCGACTCTGAGAACTTCTGAAATCGCTAGAGAATTTCTGGAATGGGATT 2280
Db 241 GlyGlnLeuSerAspSerGluLysLeuLeuGluSerLeuGluAenPheTrpAsnGlyIle 260

QY 2281 CAGGAGTGGACAGAACGACACGGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337
Db 261 GlnGluTrpThrGluArgHisGlyTyrIleValAspValSerLysArgIleProPhe 279

RESULT 10
US-09-791-537-126128
; Sequence 126128, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 133055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 126128
; LENGTH: 256
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-791-537-126128

Alignment Scores:
Pred. No.: 7.98e-121 Length: 256
Score: 1350.00 Matches: 256
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 24.5% Indels: 0
DB: 27 Gaps: 0

US-10-681-086-1 (1-3156) x US-09-791-537-126128 (1-256)

QY 1570 ATGAGATAGTGAATGGACCAATTAATGACTAGAGAAGAAAGAAATGAAGATTGTTTCAT 1629
Db 1 MetArgIleValAsnGlyProIleIleMetThrArgGluGluArgMetLysIleValHis 20

QY 1630 GAATTAAGCAACGAATATTGGATAATATGGGGATGATCTTAAGCGTATTGGTGTATTAT 1689
Db 21 GluIleLysGluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyr 40

QY 1690 GGCTCTCTTGGTCTGCTCAGACTGATGGCCCTATTTCGGATATTGAGATGATGTGTCCATG 1749
Db 41 GlySerLeuGlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMet 60

QY 1750 TCAACAGAGGAAGCAGAGTTTCAGCCATGAATGGAACACCGGTGAGTGGAGGTGGAAGTG 1809
Db 61 SerThrGluGluAlaGluPheSerHisGluIleThrThrGlyGluIleValGluVal 80

QY 1810 AATTTTTCATGACCAAGAGATTTCTACTAGATTATGATTCCTCAGGTGGATCAGATTGCGCG 1869
Db 81 AsnPheAspSerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpPro 100

QY 1870 CTTTACACATCGTCAATTTTTTCTCTATTTCGCCGATTTATGATTGAGTGGATGATCTAGAG 1929
Db 101 LeuThrHisGlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGlu 120
```

QY 1930 AAAGTGTATCAAACTGCTAAATCGGTAGAACCCCAACGTTCCACGATGCGATTGTGCC 1989  
DB 121 LysValTyrGlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAla 140  
QY 1990 CTTATCTAGAGAGCTGTTGAATATGACGAGGCAAAATGGCGTAAATATCTGTGTCAGGA 2049  
DB 141 LeuIleValGluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGly 160  
QY 2050 CCACAAACATTCTTACCATCTCTGACTGTACAGGTAGCAATGGCAGGTGCGCATGTTGATT 2109  
DB 161 ProThrThrPheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIle 180  
QY 2110 GGTCTGCATCATCGCATCTGTTATPACGACGAGCGCTTCGGTCTTAACCTGAAGCAGTTAAG 2169  
DB 181 GlyLeuHisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLys 200  
QY 2170 CAATCATGATCTCTTCAGGTATGACCATCTGTGCGCATGCTGTAATGCTGTGTCACATT 2229  
DB 201 GlnSerAspLeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeu 220  
QY 2230 TCCGACTCTGAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTTCAGGAGTGG 2289  
DB 221 SerAspSerGluLysLeuLeuSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrp 240  
QY 2290 ACAGAACGACACGGATATAGTGGATGTGTCAAAACGCATACCATTT 2337  
DB 241 ThrGluArgHisGlyTyrIleValAspValSerLysArgIleProPhe 256

RESULT 11  
US-09-791-537-80888  
; Sequence 80888, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 80888  
; LENGTH: 256  
; TYPE: PRT  
; ORGANISM: Bacillus sp  
US-09-791-537-80888

Alignment Scores:  
Pred. No.: 3.04e-120 Length: 256  
Score: 1344.00 Matches: 255  
Percent Similarity: 99.6% Conservatives: 0  
Best Local Similarity: 99.6% Mismatches: 1  
Query Match: 24.4% Indels: 0  
DB: 27 Gaps: 0

US-10-681-086-1 (1-3156) x US-09-791-537-80888 (1-256)

QY 1570 ATGAGATAGTGAATGACCAATAATATGACTAGAGAGAAAGAAATGAAGATTGTTTCAT 1629  
DB 1 MetArgIleValAsnGlyProIleLeuMetThrArgGluGluArgMetLysIleValHis 20  
QY 1630 GAAATTAAGNACGAATATGGATAATATGGGATCATGTTAAGCGTATTGGTGTAT 1689  
DB 21 GluIleLysGluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyr 40  
QY 1690 GGTCTCTGTCCTGACTGATGGCCCTATTTCGGATATTTCAGATGATGTGTGTCATG 1749  
DB 41 GlySerLeuGlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMet 60  
QY 1750 TCAACAGAGGAACGAGATTTCAGCCATGAATGGCAACCCGGTGGTGGAAAGTGG 1809

DB 61 SerThrGluGluAlaGluPheSerHisGluTrpThrThrGlyGluTrpLysValGluVal 80  
QY 1810 AATTTGATAGCAAGAGATTCTACTAGATTATGCACTCTCAGGTGGAAATCAGATTGGCCG 1869  
DB 81 AsnPheAspSerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpPro 100  
QY 1870 CTTACACATGGTCAATTTTCTCTATTTTGGCCGATTATGATTCAAGTGGATCTTAGAG 1929  
DB 101 LeuThrHisGlyGlnPhePheSerIleLeuProIleTyrAspSerGlyTyrLeuGlu 120  
QY 1930 AAAGTGTATCAAACTGCTAAATCGGTAGAACCCCAACGTTCCACGATGCGATTGTGCC 1989  
DB 121 LysValTyrGlnThrAlaLysSerValGluAlaGlnLysPheHisAspAlaIleCysAla 140  
QY 1990 CTTATCTAGAGAGCTGTTGAATATGACGAGGCAAAATGGCGTAAATATCTGTGTCAGGA 2049  
DB 141 LeuIleValGluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGly 160  
QY 2050 CCACAAACATTCTTACCATCTCTGACTGTACAGGTAGCAATGGCAGGTGCGCATGTTGATT 2109  
DB 161 ProThrThrPheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIle 180  
QY 2110 GGTCTGCATCATCGCATCTGTTATPACGACGAGCGCTTCGGTCTTAACCTGAAGCAGTTAAG 2169  
DB 181 GlyLeuHisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLys 200  
QY 2170 CAATCATGATCTCTTCAGGTATGACCATCTGTGCGCATGCTGTAATGCTGTGTCACATT 2229  
DB 201 GlnSerAspLeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeu 220  
QY 2230 TCCGACTCTGAGAACTTCTGGAATCGCTAGAGAAATTTCTGGAATGGGATTTCAGGAGTGG 2289  
DB 221 SerAspSerGluLysLeuLeuSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrp 240  
QY 2290 ACAGAACGACACGGATATAGTGGATGTGTCAAAACGCATACCATTT 2337  
DB 241 ThrGluArgHisGlyTyrIleValAspValSerLysArgIleProPhe 256

RESULT 12  
PCT-US97-07698-2  
; Sequence 2, Application PC/TUS9707698  
; GENERAL INFORMATION:  
; APPLICANT: Pachuk, Catherine J.  
; APPLICANT: Samuel, Mano  
; APPLICANT: Zurawski, John A.  
; APPLICANT: Satishchandran, C.  
; TITLE OF INVENTION: CHAIN REACTION CLONING  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris  
; STREET: One Liberty Place, 46th floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: WINDOWS  
; SOFTWARE: WordPerfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US97/07698  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/642,045  
; FILING DATE: 06-MAY-1996  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Deluca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: APOL-0294  
; TELECOMMUNICATION INFORMATION:

```
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US97-07698-2

Alignment Scores:
Pred. No.: 3,53e-119 Length: 253
Score: 1333.00 Matches: 252
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.6% Mismatches: 0
Query Match: 24.2% Indels: 0
DB: 1 Gaps: 0

US-10-681-086-1 (1-3156) x PCT-US97-07698-2 (1-253)
QY 1579 GTGAATGACCAATAATAATGACTAGAGAAAGAAATGAGATTGTTTCATGAATTAAG 1638
Db ::::|||||
1 MetAsnGlyProIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
QY 1639 GAACCAATATTGGATAAATATGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCTT 1698
Db :|||||
21 GluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyrGlySerLeu 40
QY 1699 GGTGCTCAGACTGATGGGCCCTATTTCGGATATTGAGATGATGTGTGCATGTCACACAGAG 1758
Db :|||||
41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60
QY 1759 GAACGAGAGTTCCGCCATGAATGACACACCGGTGAGTGGAGGTGGAGTGAATTTGAT 1818
Db :|||||
61 GluAlaGluPheSerHisGluTrpThrThrGlyGluTrpLysValGluValAsnPheAsp 80
QY 1819 AGCGAAGAGATTCTACTAGATTATTGATCTCAGGTGGAATCAGATTGGCCCTTACACAT 1878
Db :|||||
81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpProLeuThrHis 100
QY 1879 GGTCAATTTTCTCTATTTCGCCATTATGATTCAAGTGGATCTTAGAGAAAGTGTAT 1938
Db :|||||
101 GlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120
QY 1939 CAACCTGCTAAATCGGTAGAGCCCAACGTTCCACGATGCGATTGTGCCCTTATCGTA 1998
Db :|||||
121 GlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAlaLeuIleVal 140
QY 1999 GAAGAGCTGTTTGAATATGCAGGCAAAATGGCGTAAATATTCTGTGCAAGGACCGACAACA 2058
Db :|||||
141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
QY 2059 TTTCTACCATCCTTGACTGTACAGGTAGCAATGCGAGTGCATGTTGATTGGTCTGCAT 2118
Db :|||||
161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
QY 2119 CATGCACTCTGTTATACGACGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGAT 2178
Db :|||||
181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200
QY 2179 CTTCCTTCAGGTTATGACCATCTGTGCCAGTTCGTAATGTCTGTGTCACACTTCCGACTCT 2238
Db :|||||
201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
QY 2239 GAGAAATCTTCGGAATCGCTAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGA 2298
Db :|||||
221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
QY 2299 CACGGATATATAGTGGATGTGTCAAAACGCCATACCATTT 2337
Db :|||||
241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253
```

RESULT 13

```
PCT-US97-07853-4
; Sequence 4: Application PC/TUS9707853
; GENERAL INFORMATION:
; APPLICANT: Snyder, Linda A.
; APPLICANT: Satishchandran, C.
; TITLE OF INVENTION: CHIMERIC KANAMYCIN RESISTANCE GENE
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris
; STREET: One Liberty Place, 46th floor
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows
; SOFTWARE: WordPerfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/07853
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/642,045
; FILING DATE: 06-MAY-1996
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: APOL-0273
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 253 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US97-07853-4

Alignment Scores:
Pred. No.: 3,53e-119 Length: 253
Score: 1333.00 Matches: 252
Percent Similarity: 100.0% Conservative: 1
Best Local Similarity: 99.6% Mismatches: 0
Query Match: 24.2% Indels: 0
DB: 1 Gaps: 0

US-10-681-086-1 (1-3156) x PCT-US97-07853-4 (1-253)
QY 1579 GTGAATGACCAATAATAATGACTAGAGAAAGAAATGAGATTGTTTCATGAATTAAG 1638
Db ::::|||||
1 MetAsnGlyProIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20
QY 1639 GAACCAATATTGGATAAATATGGGATGATGTTAAGGCTATTGGTGTATTATGGCTCTCTT 1698
Db :|||||
21 GluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyrGlySerLeu 40
QY 1699 GGTGCTCAGACTGATGGGCCCTATTTCGGATATTGAGATGATGTGTGCATGTCACACAGAG 1758
Db :|||||
41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60
QY 1759 GAACGAGAGTTTCGCCATGAATGACACACCGGTGAGTGGAGGTGGAGTGAATTTGAT 1818
Db :|||||
61 GluAlaGluPheSerHisGluTrpThrThrGlyGluTrpLysValGluValAsnPheAsp 80
QY 1819 AGCGAAGAGATTCTACTAGATTATTGATCTCAGGTGGAATCAGATTGGCCCTTACACAT 1878
Db :|||||
81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpProLeuThrHis 100
QY 1879 GGTCAATTTTCTCTATTTCGCCATTATGATTCAAGTGGATCTTAGAGAAAGTGTAT 1938
Db :|||||
101 GlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120
QY 1939 CAACCTGCTAAATCGGTAGAGCCCAACGTTCCACGATGCGATTGTGCCCTTATCGTA 1998
Db :|||||
121 GlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAlaLeuIleVal 140
QY 1999 GAAGAGCTGTTTGAATATGCAGGCAAAATGGCGTAAATATTCTGTGCAAGGACCGACAACA 2058
Db :|||||
141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160
QY 2059 TTTCTACCATCCTTGACTGTACAGGTAGCAATGCGAGTGCATGTTGATTGGTCTGCAT 2118
Db :|||||
161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180
QY 2119 CATGCACTCTGTTATACGACGCGCTTCGGTCTTAACTGAAGCAGTTAAGCAATCAGAT 2178
Db :|||||
181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200
QY 2179 CTTCCTTCAGGTTATGACCATCTGTGCCAGTTCGTAATGTCTGTGTCACACTTCCGACTCT 2238
Db :|||||
201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220
QY 2239 GAGAAATCTTCGGAATCGCTAGAAATTTCTGGAATGGGATTCAGGAGTGGACAGAACGA 2298
Db :|||||
221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240
QY 2299 CACGGATATATAGTGGATGTGTCAAAACGCCATACCATTT 2337
Db :|||||
241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253
```

Db 101 GlyGlnPheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120  
QY 1939 CAAACTGCTAAATCGGTAGAGCCCAACCTTCACCATCGATTGTGCTTCATCGTA 1998  
Db 121 GlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAlaLeuIleVal 140  
QY 1999 GAAGAGCTGTTGAATATGAGGCAATGCGTAATATTCGTGTGCAAGGACCGACAACA 2058  
Db 141 GluGluLeuPheGluTyrAlaGlyLysTyrArgAsnIleArgValGlnGlyProThrThr 160  
QY 2059 TTCTACCATCCTTGACTGACGTAGAGTACAGTACGATGCGCATGCTGATGCTGCTCAT 2118  
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180  
QY 2119 CATCGCATCTGTATACGACGAGCGCTTCGTCTTAACCTGAAGCAGTAAAGCAATCAGAT 2178  
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200  
QY 2179 CTTCCTTCAGGTTATGACCATCTGTGCGCAGTTCGTAAATGCTGTCAACTTTCGCACTCT 2238  
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220  
QY 2239 GAGAACTTCGGAATCGCTAGAGAAATTCGGAATCGGATTCAGGATGAGCAGACAAGA 2298  
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTyrAsnGlyIleGlnGluTyrThrGluArg 240  
QY 2299 CACGGATATATAGTGGATGTCGTAACGCGATACCATTT 2337  
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 14

US-08-642-045A-4  
; Sequence 4: Application US/08642045A  
; GENERAL INFORMATION:  
; APPLICANT: Snyder, Linda A.  
; APPLICANT: Satishchandra, C.  
; TITLE OF INVENTION: CHIMERIC KANAMYCIN RESISTANCE GENE  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & Norris  
; STREET: One Liberty Place, 46th floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WordPerfect 6.0/6.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/642,045A  
; FILING DATE: 06-MAY-1996  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: DeLuca, Mark  
; REGISTRATION NUMBER: 33,229  
; REFERENCE/DOCKET NUMBER: APOL-0262  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 215-568-3100  
; TELEFAX: 215-568-3439  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 253 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-642-045A-4  
Alignment Scores: 3.53e-119 Length: 253  
Pred. No.: 1333.00 Matches: 252  
Score:

Percent Similarity: 100.0% Conservative: 1  
Best Local Similarity: 99.6% Mismatches: 0  
Query Match: 24.2% Indels: 0  
DB: 16 Gaps: 0  
US-10-681-086-1 (1-3156) x US-08-642-045A-4 (1-253)  
QY 1579 GTCAATGGACCAATAATAATGACTAGAGAGAAAGAAATGAAGATTGTTCAATGAATAAAG 1638  
Db 1 MetAsnGluProIleIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20  
QY 1639 GAACGAATATTGATATAATATGGGATGATGTTAAGGCTATTGGTGTGTTTATGGCTCTCTT 1698  
Db 21 GluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyrGlySerLeu 40  
QY 1699 GGTGCTCAGACTCATCGCCCTATTTCGGATATTGAGATGATGTTGTCATGTCAACAGAG 1758  
Db 41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetMetCysValMetSerThrGlu 60  
QY 1759 GAACGAGAGTTCAGCCCATGAATCGACAACCGGTGAGTGGAGGTGGAAGTGAATTTTGAT 1818  
Db 61 GluAlaGluPheSerHisGluTyrThrThrGlyGluTyrLysValGluValAsnPheAsp 80  
QY 1819 AGCAAGAGATTCTACTAGATTATGATCTCAGGTGGAATCAGATTGGCCGCTTACACAT 1878  
Db 81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTyrProLeuThrHis 100  
QY 1879 GGTCAATTTTCTCTATTGTCGATTTATGATTTCAGTTCGATCTTACAGTGAAGTGTAT 1938  
Db 101 GlyGlnPheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120  
QY 1939 CAAACTGCTAAATCGGTAGAGCCCAACCGTCCACGATCGGATTTGTGCTTCATCGTA 1998  
Db 121 GlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAlaLeuIleVal 140  
QY 1999 GAAGAGCTGTTGAATATGACGCAAAATGCGTAATATTCGTGTGCAAGGACCGACAACA 2058  
Db 141 GluGluLeuPheGluTyrAlaGlyLysTyrArgAsnIleArgValGlnGlyProThrThr 160  
QY 2059 TTCTACCATCCTTGACTGACGTAGAGTACGATGCGCATGCTGATGCTGCTCAT 2118  
Db 161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180  
QY 2119 CATCGCATCTGTATACGACGAGCGCTTCGTCTTAACTGAAGCAGTAAAGCAATCAGAT 2178  
Db 181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200  
QY 2179 CTTCCTTCAGGTTATGACCATCTGTGCGCAGTTCGTAATGCTGTCACACTTTCGACTCT 2238  
Db 201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220  
QY 2239 GAGAACTTCGGAATCGCTAGAGAAATTCGGAATCGGATTCAGGATGAGCAGACAAGA 2298  
Db 221 GluLysLeuLeuGluSerLeuGluAsnPheTyrAsnGlyIleGlnGluTyrThrGluArg 240  
QY 2299 CACGGATATATAGTGGATGTCGTAACGCGATACCATTT 2337  
Db 241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253

RESULT 15

US-09-791-537-104530  
; Sequence 104530, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Bionomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0

```
: SEQ ID NO 104530
: LENGTH: 253
: TYPE: PRT
: ORGANISM: Staphylococcus aureus
US-09-791-537-104530

Alignment Scores:
  Pred. No.:      3.53e-119      Length:      253
  Score:         1333.00        Matches:    252
  Percent Similarity: 100.0%    Conservative: 1
  Best Local Similarity: 99.6%  Mismatches:    0
  Query Match:    24.2%        Indels:      0
  DB:             27           Gaps:        0

US-10-681-086-1 (1-3156) x US-09-791-537-104530 (1-253)

QY 1579 GTGAATGACCAATAATATGACTAGACAGAAAGAAATGTTTCATGAATAATTAAG 1638
Db :|||||
1 MetAenGlyProIleMetThrArgGluGluArgMetLysIleValHisGluIleLys 20

QY 1639 GAACGATATTGGATAATATGGGATGATGTTAAGGCTATTGGTTTATGGCTCTCTT 1698
Db :|||||
21 GluArgIleLeuAspLysTyrGlyAspAspValLysAlaIleGlyValTyrGlySerLeu 40

QY 1699 GGTCTCAGACTGATGGCCCTATTTCGGATATTGAGATGATGTGTCTCATGTCAACAGAG 1758
Db :|||||
41 GlyArgGlnThrAspGlyProTyrSerAspIleGluMetCysValMetSerThrGlu 60

QY 1759 GAAGCAGAGTTTCAGCCATGAATGGACAACCGGTGAGTGGAGGTGGAAAGTGAATTTTGAT 1818
Db :|||||
61 GluAlaGluPheSerHisGluTrpThrThrGlyGluTrpLysValGluValAsnPheAsp 80

QY 1819 AGCGAAGAGATTTCTACTAGATTATGCTCAGGTGGATCAGATTGGCGCTTACACAT 1878
Db :|||||
81 SerGluGluIleLeuLeuAspTyrAlaSerGlnValGluSerAspTrpProLeuThrHis 100

QY 1879 GGTCAATTTTCTCTATTTCGCGATTATGATTCAGGTGGATACCTTAGAGAAAGTGTAT 1938
Db :|||||
101 GlyGlnPhePheSerIleLeuProIleTyrAspSerGlyGlyTyrLeuGluLysValTyr 120

QY 1939 CAAACTGCTAAATCGGTAGAGCCCAACGTTCCACGATGCGATTTGTGCCCTTATCGTA 1998
Db :|||||
121 GlnThrAlaLysSerValGluAlaGlnThrPheHisAspAlaIleCysAlaLeuIleVal 140

QY 1999 GAAGAGCTTTTGATATGACGGCAAAATGGCGTAATATTCGTGTGCAAGGACCGACACA 2058
Db :|||||
141 GluGluLeuPheGluTyrAlaGlyLysTrpArgAsnIleArgValGlnGlyProThrThr 160

QY 2059 TTCTACCATCTTCTGACTGTACAGGTAGCAATGGCAGGTGCCATGTTGATTGGTCTGCAT 2118
Db :|||||
161 PheLeuProSerLeuThrValGlnValAlaMetAlaGlyAlaMetLeuIleGlyLeuHis 180

QY 2119 CATCGCATCTGTTATACGACGAGCCGCTTCGGTCTTAACTGAAGCAGTTTAAGCAATCAGAT 2178
Db :|||||
181 HisArgIleCysTyrThrThrSerAlaSerValLeuThrGluAlaValLysGlnSerAsp 200

QY 2179 CTTCCTTCAGGTATGACCATCTGCGCAGTTCGTAATGCTCGTCAACTTTCGACTCT 2238
Db :|||||
201 LeuProSerGlyTyrAspHisLeuCysGlnPheValMetSerGlyGlnLeuSerAspSer 220

QY 2239 GAGAAACTTCTGGAAATCCGCTAGAGAAATTTCTGGAATGGGATTTCAGGAGTGGACAGAACGA 2298
Db :|||||
221 GluLysLeuLeuGluSerLeuGluAsnPheTrpAsnGlyIleGlnGluTrpThrGluArg 240

QY 2299 CACGATATATAGTGGATGTGTCAAAACGCATACCATTT 2337
Db :|||||
241 HisGlyTyrIleValAspValSerLysArgIleProPhe 253
```

Search completed: February 10, 2006, 03:48:56  
Job time : 1376 secs



Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	1836	33.3	389	8	US-60-752-355-27114	Sequence 27114, A
2	1676	30.4	371	8	US-60-752-355-47667	Sequence 47667, A
3	942	17.1	387	8	US-60-752-355-46281	Sequence 46281, A
4	859	15.6	388	8	US-60-752-355-9395	Sequence 9395, A
5	843	15.3	392	8	US-60-752-355-26171	Sequence 26171, A
6	804	14.6	391	8	US-60-752-355-17511	Sequence 17511, A
7	762	13.8	390	8	US-60-752-355-25521	Sequence 25521, A
8	743	13.5	400	8	US-60-752-355-21930	Sequence 21930, A
9	721	13.1	393	8	US-60-752-355-36490	Sequence 36490, A

```
Db 21 ArgAsnLeuArgSerMetAspGlyAlaProValProGluArgAsnIleAspGlyGluAsn 40
Qy 142 CAAACGGCTCTGCTCTCAAAACAATATTTAGCGCTCGCAAGCATAGAGCTTTGATCGAT 201
Db 41 GlnThrValTrpSerSerAsnAsnTrpLeuGlyLeuAlaSerAspArgArgIleAsp 60
Qy 202 CGAGCCCAACAGCATTTGAGCAATTTGGCAGGAGCAGCGGTTTCAGCTTTAAACGACA 261
Db 61 AlaAlaGlnThrAlaLeuGlnGlnPheGlyThrGlySerSerGlySerArgLeuThrThr 80
Qy 262 GCGAATTCGGCTCGCATGAAAGCTAGAAAGAGATTCGAGCTTTAAACTGACAGAA 321
Db 81 GlyAsnSerValTrpPheGluLysLeuGluLysLysIleAlaSerPheLysLeuThrGlu 100
Qy 322 GCGGCCCTGCTGCTTTTCAGCGGTTTACTTTGGCAATGTCGGTGCTCTTTTCATCCTTGCCA 381
Db 101 AlaAlaLeuLeuPheSerSerGlyTrpLeuAlaAsnValGlyValLeuSerSerLeuPro 120
Qy 382 GAAAGGAGATGTCTATTTAAGTGACCAGCTCAATCATGCAAGTATGATCGACGGCTGC 441
Db 121 GluLysGluAspValIleLeuSerAspGlnLeuAsnHisAlaSerMetIleAspGlyCys 140
Qy 442 CGACTTCTAAGCGTGATACAGTGTGTTATCGGCATATTTGATGATGATGATGATGATGAT 501
Db 141 ArgLeuSerLysAlaAspThrValValTrpArgHisIleAspMetAsnAspLeuGluAsn 160
Qy 502 AAGCTGAATGAACACAGCGTTATCAGCGCGGTTTATCGTAACAGCGGAGTATTCAGC 561
Db 161 LysLeuAsnGlnThrGlnArgTrpGlnArgPheIleValThrAspGlyValPheSer 180
Qy 562 ATGGATGGCACAAATCGCCCTCTTGATCAGATCATCTCACTTGGCAACGGTATCATGCC 621
Db 181 MetAspGlyThrIleAlaProLeuAspGlnIleSerLeuAlaLysArgTrpHisAla 200
Qy 622 TTCGTGCTGCTGATGATGCCCAACAGAGAGTTTGGCGATTCGGGCGATTCGGGCAAGGACG 681
Db 201 PheValValValAspAspAlaHisAlaThrGlyValLeuGlyAspSerGlyGlnGlyThr 220
Qy 682 AGTGAATCTTTGGTGTTCGCCGACATTTGATCGGCACCTTAAGCAAGCTGTGGC 741
Db 221 SerGluTrpPheGlyValCysProAspIleValIleGlyThrLeuSerLysAlaValGly 240
Qy 742 GCGGAAGAGGTTTTCGCGCAGGATTCAGCGGTCTTCATCGACTTTTCTGCTGAACCATGCC 801
Db 241 AlaGluGlyGlyPheAlaAlaGlySerAlaValPheIleAspPheLeuLeuAsnHisAla 260
Qy 802 AGAATATTTCTTTCAACCGCTATTCGCCAGCCAGCTGTGGCGGTCTTCAGAGGCT 861
Db 261 ArgThrPheIlePheGlnThrAlaIleProProAlaSerCysAlaAlaHisGluAla 280
Qy 862 TTCACATCATTTGAAGCCAGCAGGAAACAGCAGCTTTTATTTCTTATATCAGCATG 921
Db 281 PheAsnIleIleGluAlaSerArgGluLysArgGlnLeuLeuPheSerTrpIleSerMet 300
Qy 922 ATCAGAACCACTGCAAGAATATGGGTATGTGTGAAAGGAGATTCACACACCGATTAT 981
Db 301 IleArgThrSerLeuLysAsnMetGlyTrpValValLysGlyAspHisThrProIleIle 320
Qy 982 CCTGTAGTCATTTGGCGATGCCCATAAACGGTCTCTATTTGCTGAAAACTCAGGGCAAG 1041
Db 321 ProValValIleGlyAspAlaHisLysThrValLeuPheAlaGluLysLeuGlnGlyLys 340
Qy 1042 GGAATTTATGCTCTGCTGCTGCGCCCAACCGTTCGCGCGGTGAAGCCGATTCGA 1101
Db 341 GlyIleTrpAlaProAlaIleArgProProThrValAlaProGlyGluSerArgIleArg 360
```

## RESULT 2

US-60-752-355-47667

; Sequence 47667, Application US/60752355

; GENERAL INFORMATION:

; APPLICANT: Abad, Mark et al.

; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits

; FILE REFERENCE: 38-21(53720)B

; CURRENT APPLICATION NUMBER: US/60/752,355

; CURRENT FILING DATE: 2005-12-21

; NUMBER OF SEQ ID NOS: 52803

; SOFTWARE: PatentIn version 3.3

; SEQ ID NO 47667

; LENGTH: 371

; TYPE: PRN

; ORGANISM: Bacillus subtilis

US-60-752-355-47667

## Alignment Scores:

Pred. No.:	7,04e-157	Length:	371
Score:	1676.00	Matches:	330
Percent Similarity:	93.3%	Conservative:	6
Best Local Similarity:	91.7%	Mismatches:	6
Query Match:	30.4%	Indels:	18
DB:	8	Gaps:	1

US-10-681-086-1 (1-3156) x US-60-752-355-47667 (1-371)

```
Qy 22 TTGAGATTTGATTTCTCTGTTTAAACGACGGTTAGACAGAATGAAGAAGCGCGGTACAT 81
Db 1 MetLysIleAspSerTrpLeuAsnAspArgLeuAspIleAlaLysGluAlaGlyValHis 20
Qy 82 CGTAACCTCGCGTCAATGGATGGAGCCCGGTTCCAGAGAGGAATATTGATGGCAAAAT 141
Db 21 ArgAsnLeuArgSerMetAsnGlyAlaProValProGluArgAsnIleAspGlyGluAsn 40
Qy 142 CAAACGGTCTGGTCTCTCAAAACAATTTATTAGGGCTCCAGCGCATAGACGTTTGCATCGAT 201
Db 41 GlnThrValTrpSerSerAsnAsnTrpLeuGlyLeuAlaSerAspArgGLeuIleAsp 60
Qy 202 GCAGCCCAACAGCATTTGACCAATTTGGACAGGAGCAGCGGTTACGCTTTTAACGACA 261
Db 61 AlaAlaGlnThrAlaLeuGlnGlnPheGlyThrGlySerSerGlySerArgLeuThrThr 80
Qy 262 GGCATTTCTGCTGCGCATGAAAGCTAGAAAGAGATTCGCCAGCTTTAAACTGACAGAA 321
Db 81 GlyAsnSerValTrpHisGluLysLeuGluLysLysIleAlaSerPheLysArgThrGlu 100
Qy 322 GCGGCCCTGCTGTTTTCGACGGTACTTTCGCCAATGTTCGGTGTCTTTCATCTCTGCCA 381
Db 101 AlaAlaLeuLeuPheSerSerGlyTrpLeuAlaAsnValGlyValLeuSerSerLeuPro 120
Qy 382 GAAAGGAGATGCATTTTAAAGTGACCGACTCAATCATGCAAGTATGATCGACGGCTGC 441
Db 121 GluLysGluAspValIleLeuSerAspGlnLeuAsnHisAlaSerIleIleAspGlyCys 140
Qy 442 CGACTTCTAAGCGTGATACAGTGTGTTTATCGGCATATTCATATGATGATGATGATGATGAT 501
Db 141 ArgLeuSerLysAlaAspThrValValTrpArgHisIleAspMetAsnAspLeuGluAsn 160
Qy 502 AAGCTGAATGAAACACAGCGTTTATCAGCGCGGTTTATCGTAAACAGCGGATTTTCAGC 561
Db 161 LysLeuAsnGlnThrGlnArgTrpGlnArgPheIleValThrAspGlyValPheSer 180
Qy 562 ATGATGGCACAATCGCCCTCTTGATCAGATCATCTCACTTGGCAACCGTATCATGCC 621
Db 181 MetAspGlyThrIleValProLeuAspGlnIleSerLeuAlaLysArgTrpHisAla 200
Qy 622 TTCGTGCTGCTGATGATGCCCAACAGGAGTTTGGCGATTCGGGCAAGGAAACG 681
Db 201 PheValValValAspAspAlaHisAlaThrGlyValLeuGlyAspSerGlyArgGlyThr 220
Qy 682 AGTGAATCTTTGGTGTTCGCCGACATTTGATCGGCACCTTAAGCAAGCTGTGGC 741
Db 221 SerGluTrpPheGlyValTrpProAspIleValIleGlyThrLeuSerLysAlaValGly 240
Qy 742 GCGGAAGAGGTTTTCGCGCAGGATTCAGCGGTCTTCATCGACTTTTCTGCTGAACCATGCC 801
Db 241 ThrGluGlyGlyPheAlaAlaGlySerAlaValPheIleAspPheLeuLeuAsnHisAla 260
Qy 802 AGAACATTTTATCTTTCAACCGCTATTCGCCAGCCAGCTGTGGCGGTCTTCAGAGGCT 861
Db 802 AGAACATTTTATCTTTCAACCGCTATTCGCCAGCCAGCTGTGGCGGTCTTCAGAGGCT 861
```

```
Db 261 ArgThrPheIlePheGlnThrAlaIleProProAlaSerCysAlaAlaAlaHisGluAla 280
QY 862 TTCACATCATTCAGCCAGCAGGGAACAGCAGCTTTTATTTCTTATATCAGCATG 921
Db 281 PheAsn-----282
QY 922 ATCAGAACCATCTGAAGATATGGGTATGTGGTGAAGAGAGATCACACACCGATTATT 981
Db 283 IleArgThrSerLeuLysAsnMetGlyTyrValValLysGlyAspHisThrProIleIle 302
QY 982 CTGTAGTCAATGGCGATGCCATATAAACCAGTCTATTCTGCTGAAACAACTGCAGGGCAAG 1041
Db 303 ProValValIleGlyAspAlaHisLysThrValIlePheAlaGluLysLeuGlnGlyLys 322
QY 1042 GGAATTATGCTCTCCATTCGGCCGCAACCGTTGCGCGCGGTGAAAGCCGATTCGA 1101
Db 323 GlyIleTyrAlaProAlaIleArgProProThrValAlaProGlyGluSerArgIleArg 342
RESULT 3
US-60-752-355-46281
; Sequence 46281, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 46281
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Kurthia sp. 538-KA26
US-60-752-355-46281
Alignment Scores:
Pred. No.: 3,11e-84 Length: 387
Score: 942.00 Matches: 184
Percent Similarity: 73.6% Conservative: 81
Best Local Similarity: 51.1% Mismatches: 87
Query Match: 17.1% Indels: 8
DB: 8 Gaps: 4
US-10-681-086-1 (1-3156) x US-60-752-355-46281 (1-387)
QY 37 TGGTTAAACGAGCGGTTAGACAGATGAAGAGCCGCGCTACATCGTAACCTCGGTCA 96
Db 3 TrpGluLysGlu---LeuGluLysIleLysGluGlyLeuTyrArgGlnLeuGlnThr 21
QY 97 ATGATGCGAGCGCGGTTCCAGAGAGGAAT-----ATTGATGCGGAAATCAAACG 147
Db 22 ValGlu-----ThrMetSerAspGlnGlyTyrAlaMetValAsnGlyLysLysMetMet 39
QY 148 GTCTGTGCTCAACAATATTATAGGCTCGCAAGCATAGACGTTTGTATCGATGCGAGCC 207
Db 40 MetPheAlaSerAsnIleTyrLeuGlyIleAlaAsnAspGlnArgLeuIleGluAlaSer 59
QY 208 CAACACGATTCGACGAATTTGGACAGAGAGCGGTTCCAGCTTTTAAACGACAGGCAAT 267
Db 60 ValGlnAlaThrGlnArgPheGlyThrGlySerThrGlySerArgLeuThrThrGlyAsn 79
QY 268 TCGTCTGGCATGAAGCTAGAAAGACAGATTCAGCTTAACTGACAGAGCGCGCC 327
Db 80 ThrIleValHisGluLysLeuGluLysArgLeuAlaGluPheLysGlnThrAspAlaAla 99
QY 328 CTGCTGTTTTCGAGCGGTACTTCGGCAATGTCGGTGTCTCTTTCATCTTCCTTCGCCAGAAAAG 387
Db 100 IleValLeuAsnThrGlyTyrMetAlaAsnIleAlaAlaLeuThrThrLeuValGlySer 119
QY 388 GAAGATGCTATTTAAGTACCGACTCAATCATGCAAGTATGATCGAGCGGTCGCCGACTT 447
Db 120 AspAspLeuIleLeuSerAspGluMetAsnHisAlaSerIleIleAspGlyCysArgLeu 139
```

```
QY 448 TCTAAGGCTGATACAGTTGTTTATCGGCATATTTGATATGATATCATCTTGAACAAAGCTG 507
Db 140 SerArgAlaGluThrIleTyrArgHisAlaAspLeuLeuAspLeuGluMetLysLeu 159
QY 508 AATGAACACACAGGTTATCAGCCGCTTTTATCGTAACAGAGCGGATTTACGATGGAT 567
Db 160 GlnIleAsnThrArgTyrArgLysArgIleIleValThrAspGlyValPheSerMetAsp 179
QY 568 GGCACAAATCGCCCTTCGATCAGATCATCTCACITTCGCAACCGCTATCATGCTTCGTG 627
Db 180 GlyAspIleAlaProLeuProGlyIleValGluLeuAlaLysArgTyrAspAlaLeuVal 199
QY 628 TCGTGTATGATGCCCAACAGAGTTTGGCGCATTCGGGCAACAGGAAACAGAGTGAA 687
Db 200 MetValAspAspAlaHisAlaThrGlyValLeuGlyLysAspGlyArgGlyThrSerGlu 219
QY 688 TACTTTGGTGTGTTGCTCC-----GACATGTATTTCGGCACCTTAAGCAAGCTGTGGC 741
Db 220 HisPheGlyLeuLysGlyLysIleAspIleGluMetGlyThrLeuSerLysAlaValGly 239
QY 742 GCGGAAGGAGGTTTTCGCGCAGGATCAGCGGTCTTCATCGACTTTTGTCTGAACCATGCC 801
Db 240 AlaGluGlyGlyTyrIleAlaGlySerArgSerLeuValAspTyrValLeuAsnArgAla 259
QY 802 AGAACATTTATCTTTCAAACCGCTATTCCGCCAGCAGCTGTGCGGCTGCTCAGAGGCT 861
Db 260 ArgProPheValPheSerThrAlaLeuSerAlaGlyValValAlaSerAlaLeuThrAla 279
QY 862 TTCACATCATTCAGCCAGCAGGGAACAAACGACACTTTTATTTCTTATATCAGCATG 921
Db 280 ValAspIleIleGlnSerGluProGluArgValArgIleArgAlaMetSerGlnArg 299
QY 922 ATCAGAACCATCTGAAGATATGCGTATGTGTAAGAGGATCACACACCGATTATT 981
Db 300 LeuTyrAsnGluLeuThrSerLeuGlyTyrThrValSerGlyGlyGluThrProIleLeu 319
QY 982 CTGTAGTCAATGGCGATGCCATAAAACGGTCTATTGCTGAAACAACTGCAGGGCAAG 1041
Db 320 AlaIleIleCysGlyGluProGluGlnAlaMetPheLeuSerLysGluLeuHisLysHis 339
QY 1042 GGAATTTATGCTCTCCATTCGGCGCCGCAACCGTTGCGCGCGGTGAAAGCCGATTCGA 1101
Db 340 GlyIleTyrAlaProAlaIleArgSerProThrValProLeuGlyThrSerArgIleArg 359
RESULT 4
US-60-752-355-9295
; Sequence 9295, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 9295
; LENGTH: 388
; TYPE: PRT
; ORGANISM: Clostridium thermocellum ATCC 27405
US-60-752-355-9295
Alignment Scores:
Pred. No.: 4,54e-76 Length: 388
Score: 859.50 Matches: 181
Percent Similarity: 65.5% Conservative: 60
Best Local Similarity: 49.2% Mismatches: 118
Query Match: 15.6% Indels: 9
DB: 8 Gaps: 4
US-10-681-086-1 (1-3156) x US-60-752-355-9295 (1-388)
QY 52 TTAGACAGAAATGAAGAGCCGCGGTACATCGTAACCTCGGTCAATGGATGGAGCGCG 111
```

```
Db      8 LeuLysAspIleLysAsnLysGlyLeuTyrArgGluPheArgAsnValAsnAlaGln 27
Qy      112 GTTCAGAGAGGAATATTGATGGCGGAATCAAAACGGTCTGGTCTCAAAACAATTATTTA 171
Db      28 GlyProTyrThrValIleAspGlyArgLysMetLeuMetSerAsnAsnTyrLeu 47
Qy      172 GGGTCGCAAGCGATAGACGTTTGATCGATGAGCCCAACAGCATTCAGCAATTTGGG 231
Db      48 GlyLeuCysAspAspIleArgLeuLysArgAlaIleGluSerIleArgLysPheGly 67
Qy      222 ACAGAAACGAGGTTTCAGTTTAAACGACAGCAATTCGGTCTGGCATGAAGAGTAGAA 291
Db      68 ValGlyAlaGlyGlySerArgLeuThrCysGlyAsnPheGluLeuHieGluLeuGlu 87
Qy      292 AAGAAGATTGCCAGCTTTAAACTGACAGAGCGGCCCTGCTGTTTCAGCGGTACTTG 351
Db      88 GluArgLeuAlaLysPheLysAspValGluSerCysIleValPheGlySerGlyTyrAla 107
Qy      352 GCCAATGTCGGTGTCTTTTCATCTTCGACAGAAAGAGATGTCATTTAAAGTGACGAG 411
Db      108 AlaAsnIleGlyAlaIleSerGlyIleAlaAspLysAsnTrpValIlePheCysAspArg 127
Qy      412 CTCATCATGCAAGTATGATCGAGCGCTGCCGACTTCTTAAGCGTGATPACAGTTGTTAT 471
Db      128 LeuAsnHieAlaSerIleValAspGlyIleArgLeuSerGlyAlaLysLeuValValTyr 147
Qy      472 CGGCATATTGATCAATGATCTTGAAACAAAGCTGAATGAACACACGCTTATCAG--- 528
Db      148 LysHisCysAspMetGluAspLeuGluSerLysIle-----ValArgTyrHisThr 164
Qy      529 ---GCCCGCTTTATCGTAACAGACGGAGTATTACGATGATGGCACATCGCCCTCTT 585
Db      165 GlyLysSerLeuIleValThrAspGlyValPheSerMetAspGlyAspValAlaProVal 184
Qy      586 GATCAGATCATCTCATTGCGAAACGCTATCATGCCTTCGTCGTCGTTGATGATGCCAC 645
Db      185 AspArgIleValLysLeuAlaLysLysTyrAsnLeuMetThrMetValAspAspAlaHis 204
Qy      646 GCACAGAGATTGCGGCGATTCGGGCAACGAGCAAGAGAGTGAATCTTTGGTGT--- 699
Db      205 AlaThrGlyIleLeuGlyGlyLysGlyArgGlyThrSerGluTyrPheGlyLeuLysAsp 224
Qy      700 TGTCCGACATGTTATCGGACCTTAAGCAAAAGCTGTTCGCGGAGGAGGTTTTCG 759
Db      225 AlaValAspIleSerMetGlyThrLeuSerLysAlaPheGlyValGluGlyGlyPheVal 244
Qy      760 GCAGATCAGCGGTCTTCATCGACTTTTGTGTAACCATCGCAGAACATTTATCTTCAA 819
Db      245 AlaGlyLysArgLysLeuValAspPheLeuArgHisLysAlaLysSerPheIleTyrSer 264
Qy      820 ACCGCTATTCGCCAGCCAGCTGTGCGCTGCTCAGCGGCTTTCACATCATTTGAAGCC 879
Db      265 ThrAlaProProHisAsnMetAlaAlaAlaLeuGluAlaLeuAsnIleIleGluThr 284
Qy      880 AGCAGGAAAAACGACAGCTTTATTTCTTATATATATATATATATATATATATATAT 939
Db      285 GluProGlnAlaArgLysGluLeuAlaGluLysSerValTrpLeuArgAsnArgLeuIle 304
Qy      940 AATATGGGTTATGTGGTGAAGAGATCACACCCGATTAATTCCTGTAGTCAATTTGGCGAT 999
Db      305 GluLysGlyPheAsnValProLysGlyValThrProIleIleProLeuMetValGlyAsp 324
Qy      1000 GCCCATAAACGGTCTTATTTGCTGAAACAACTGAGCGGCAAGGAATTTATGCTCTCGCC 1059
Db      325 ValAsnThrAlaValGluPheSerMetLeuLeuTyrAsnGluIleTyrIleProAla 344
Qy      1060 ATTGGCGCGCAACCGTTGCGCGCGGTGAAGCGGATTCGA-----AGCTTCGGCGAGC 1113
Db      345 IleArgProProThrValProLysGlyThrSerArgLeuArgIleSerIleMetAlaSer 364
Qy      1114 AGGTCGAGATCAGGAATGAGTTT 1137
```

```
Db      365 HisSerTyrGluAspMetGluPhe 372
RESULT 5
US-60-752-355-26171
; Sequence 26171, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)/B
; CURRENT APPLICATION NUMBER: US/60752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 26171
; LENGTH: 392
; TYPE: PRT
; ORGANISM: Bacillus subtilis subsp. subtilis str. 168
US-60-752-355-26171
Alignment Scores:
Pred. No.: 1,96e-74 Length: 392
Score: 843.00 Matches: 169
Percent Similarity: 66.8% Conservative: 77
Best Local Similarity: 45.9% Mismatches: 120
Query Match: 15.3% Indels: 2
DB: 8 Gaps: 1
```

US-10-681-086-1 (1-3156) x US-60-752-355-26171 (1-392)

```
Qy      37 TGGTTAAACGAGCGGTTAGACAGAAAGAAAGACCGCGGTACATCGTAACCTCGGTCA 96
Db      7 PheLeuLysAlaGluLeuAsnSerMetLysGluAsnHisThrTrpGlnAspIleLysGln 26
Qy      97 ATGATGCGAGCGCGGTTCCAGAGAGGAATATTGATCGCGAANAATCAACGGTCTGGTCC 156
Db      27 LeuGluSerMetGlnGlyProSerValThrValAsnHisGlnLysValIleGlnLeuSer 46
Qy      157 TCAACAATATTATTAGGCGTCGCAAGGATAGAGTTGATTCGATCGACGCCCAACAGCA 216
Db      47 SerAsnAsnTyrLeuGlyPheThrSerHisProArgLeuIleAsnAlaIleGlnGluAla 66
Qy      217 TTCAGCAATTTGGGACAGGAGCGGTTTCACGTTTACAGCAGCAGCAATTCGGTCTGG 276
Db      67 ValGlnGlnTyrGlyAlaGlyThrGlySerValArgThrIleAlaGlyThrPheThrMet 86
Qy      277 CATGAAAAGCTAGAAAAGAGATTCGCAGCTTTTAACTGACAGAGCGCCCTCTGTTT 336
Db      87 HieGlnGluLeuGluLysLysLeuAlaAlaPheLysLysThrGluAlaAlaLeuValPhe 106
Qy      337 TCAGAGCGGTTACTTGGCCAATGTCGGTGTCTCTTTTCATCTTCCAGAAAAGGAAGATGTC 396
Db      107 GluSerGlyPheThrThrAsnGlnGlyValLeuSerSerIleLeuSerLysGluAspIle 126
Qy      397 ATTTTAAGTACGACGCTCAATCATGCAAGTATGATCGACGCGCTGCCGACTTTTAAGGCT 456
Db      127 ValIleSerAspGluLeuAsnHisAlaSerIleIleAspGlyIleArgLeuThrLysAla 146
Qy      457 GATACAGTTGTTTATCGGCATATTGATGATGATCTTCAAAACAGCTGAATGAACA 516
Db      147 AspLysLysValTyrGlnHisValAsnMetSerAspLeuGluArgValLeuArgLysSer 166
Qy      517 CAGCGTTTATCAGCGCGCTTTTTCGTAACAGAGCGAGTATTTCAGCATGGATGGCACAACT 576
Db      167 MetAsnTyrArgMetArgLeuIleValThrAspGlyValPheSerMetAspGlyAsnIle 186
Qy      577 GCCCTCTTTCATCAGATCATCTCATTGCGAAACGCTATCATGCTTCGTCGTCGTCGTCAT 636
Db      187 AlaProLeuProAspIleValGluLeuAlaGluLysTyrAspAlaPheValMetValAsp 206
Qy      637 GATGCCCAACCAACAGAGAGTTTGGCGGATTCGGGACAGAGACGAGTGAATACTTTGGT 696
Db      207 AspAlaHisAlaSerGlyValLeuGlyGluAsnGlyArgGlyThrValAsnHisPheGly 226
```



```
; SEQ ID NO 25521
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Exiguobacterium sp. 255-15
US-60-752-355-25521

Alignment Scores:
  2,03e-66      Length: 390
  762.00       Matches: 157
  62.6%        Conservative: 67
  43.9%        Mismatches: 132
  13.8%        Indels: 2
  8            Gaps: 1
DB:

US-10-681-086-1 (1-3156) x US-60-752-355-25521 (1-390)
Qy 40 TTAACGAGCGGTTAGACAGATGAAGAGCGCGTACATCGTAACTCGCGTCAATG 99
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 6 LeuArgThrGluLeuGluMetLysGlnAlaGlyThrPheArgGluLeuValAlaLeu 25
Qy 100 GATGAGCGCGGTTCCAGAGGAGATATTGATGCGGAATCAACCGTCTGCTCTCA 159
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 26 GluSerAlaGlnHisAsnArgValThrValAspGlyLysGluLeuGlnLeuSer 45
Qy 160 AACAAATTATTTAGGCTCGCAAGCATAGACGTTTTCATCGATGCGCCCAACAGCAT 219
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 46 AsnAsnTyrLeuGlyLeuAlaHisProArgLeuAlaLysArgAlaAspAlaLeu 65
Qy 220 CAGCAATTGGGACAGGAGCGGTTCCAGTTTAAACGACAGGCAATTCGGTCTGGCAT 279
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 66 LeuGluPheGlyAlaGlyThrGlySerValArgThrIleAlaGlyThrLeuGluMeth 85
Qy 280 GAAAAGCTAGAAAGAGATGCCAGCTTTAACTGACAGAGCGGCTCTCTGTTTCG 339
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 86 GlnAlaPheGluArgGluLeuAlaThrPheLysHisThrGluAlaAlaLeuValPhe 105
Qy 340 AGCGGTTACTTGGCCAAATGTCGGTCTCTTCATCTTCCAGAAAGAAAGATGTCATT 399
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 106 SerGlyPheAlaThrAsnLeuGlyValLeuSerAlaLeuLeuGlyProGluAspVal 125
Qy 400 TTAAGTACCAGCTCAATCATGCAAGTATGATCGACGCTCCGCACTTCTTAAGCTGAT 459
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 126 IleSerAspAlaLeuAsnHisAlaSerIleIleAspGlyIleArgLeuThrLysAla 145
Qy 460 ACAGTTGTTTATCGCATATTGATATGAATGATCTTGAACCAAGCTGAATGAACACAG 519
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 146 ArgArgIleTyrAsnHisValAspLeuAlaAspLeuGluAlaLeuLeuGlnThrGln 165
Qy 520 GTTATACGCGCGGTTTATCGTACACAGCGAGTATTACGATGGATGGCACATCGCC 579
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 166 AspAlaArgThrArgLeuValValThrAspGlyValPheSerMetAspGlyAsnIleAla 185
Qy 580 CCTCTTGATCATCATCTCACTTGCAGAACGCTATCATGCTTCGTGCTGCTGATGAT 639
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 186 ProLeuProGluIleValGluLeuAlaGluLysTyrAspAlaLeuValMetValAsp 205
Qy 640 GCCACGCAACAGAGTTTGGCGCATTCGGGACAGGAGCAAGCAAGTGAATCTTGGTGT 699
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 206 AlaHisAlaSerGlyValLeuGlyLysSerGlyArgGlyThrValAsnHisPheGlyLeu 225
Qy 700 TGTCCGACATGTT-----ATCGGCACCTTAAGCAAGCTGTGGCGCGGAGGAGGT 753
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 226 AspGlyArgValAlaLeuGlnValGlyThrLeuSerLysAlaIleGlyValLeuGlyGly 245
Qy 754 TTGCGGAGGATCAGCGGCTTTCATCGACTTTTGTGTAACCATGCGCAACATTTATC 813
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 246 TyrValAlaCysGluGlnHisValLysAspTyrLeuIleHisLysGlyArgProPheLeu 265
Qy 814 TTTCAAACCGCTATTCGCCGACCGCATGTCGGGCTGCTCAGAGGCTTTCACATCAT 873
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 266 PheSerThrHisProProAlaValValGluAlaAsnArgGluAlaLeuArgValMet 285
Qy 874 GAAGCCAGCAGGAGAAACGACAGCTTTTATTTTCTTATATCAGCATGATCAGAACCAT 933
```

```
Db 286 GluGluGluThrGluLeuPheAspArgLeuTrpGluAsnThrGluPheLysHisGly 305
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Qy 934 CTGAAGAATATGGGTTATGGTGAAGAGAGATCACACACCGATATTCTGTAGTCATT 993
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 306 LeuArgGluLeuGlyPheAsnIleGlyThrSerThrThrProIleThrProValIleVal 325
Qy 994 GGCGATGCCCATAAACCGGCTCTATTGCTGAAACCTGACAGGCAAGGGAATTTATGCT 1053
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 326 GlyAspGluAlaLeuCysHisGlnLeuSerAspArgLeuArgGlnHisGlyValPheAla 345
Qy 1054 CCTGCCATTGCGCGCCCAACCGTTGCGCGCGTGAAGCCGATTCGAAGCTTG 1107
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 346 GlnGlyIleAlaPheProThrValAlaLysGlyLysAlaArgValArgThrIle 363

RESULT 8
US-60-752-355-21930
; Sequence 21930, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 21930
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Haloarcula marismortui ATCC 43049
US-60-752-355-21930

Alignment Scores:
  1.39e-64      Length: 400
  743.50       Matches: 161
  62.3%        Conservative: 69
  43.6%        Mismatches: 120
  13.5%        Indels: 19
  8            Gaps: 6
DB:

US-10-681-086-1 (1-3156) x US-60-752-355-21930 (1-400)
Qy 40 TTAACGAGCGGTTAGACAGATGAAGAGCGCGTACATCGTAACTCGCGTCAATG 99
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 7 LeuAsnGluArgLeuGluGlnArgAspAlaGlnAsnLeuArgHisLeuGluValAla 26
Qy 100 GATGAGCGCGGTTCCAGAGAGGAATATTGAT-----132
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 27 GluSerValSerAlaArgThrArgPheAlaAspProLysGlyAspProGluPhe 46
Qy 133 GGCGAAAATCAACGGTCTGGTCTCTCAACCAATTTATTAGGCTCGCAAGGATAGACGT 192
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 47 GlyAspGluGluValValPheAlaAlaAsnAsnTyrLeuGlyLeuAlaAspSerArg 66
Qy 193 TTGATCATGATCGACCCCAACAGCATTCGCAATTTGGGACAGGAGCGGTTCAAGT 252
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 67 ValGlnArgAlaAlaGluLeuGlyAlaArgThrValGlySerGlyAlaGlyAlaSerArg 86
Qy 253 TTAACGACAGCGCAATTCGGTCTGCGCATGAAAGCTAGAAAGAGATTGCCAGCTTTAAA 312
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 87 LeuValThrGlyAspThrLysValHisArgAlaLeuGluArgAspLeuAlaSerLys 106
Qy 313 CTGACAGAGCGGCTCTGTTTCGAGCGGTTACTTGGCCATGTCGGTGTCTCTTCA 372
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 107 GlyThrGluArgAlaLeuValPheSerSerGlyTyrAlaAlaAsnIleGlyThrIleAsp 126
Qy 373 TCCTTGGCAGAAAAGGAAGATGTCATTTAAGTCACCACTCAATCATGCAAGTATGATC 432
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 127 AlaLeu-----AlaProAspValValPheSerAspGluLeuAsnHisAlaSerIleIle 144
Qy 433 GACGGCTCGGCACTTTCTTAAGGCTGTACAGCTGTTTATCGGCATATTGATATGAATGAT 492
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 145 AspGlyCysArgValGlyAlaSerGluThrValValTyrAspHisCysAspProAspAsp 164
```

```
QY 493 CTTGAAACAAAGCTG-----AATGAAACACAGCGTTATCAGCGCCGTTTATC 540
    |||
    |||
    |||
Db 165 LeuArgAlaLysLeuAlaThrArgAlaAlaAspValAspGluAlaGlnTrpLeuVal 184
    :
    :
    :
QY 541 GTAACAGACGGATTTACAGATGATGGACAAATCGCCCTCTTGATCAGATCATCTCA 600
    |||
    |||
    |||
Db 185 ValThrAspSerValPheSerMetAspGlyAspIleAlaProLeuSerAlaIleCysAsp 204
    |||
    |||
    |||
QY 601 CTTGCGAAACGCTATCATGCTCTGCTGCTGATGATGCCACGCAACAGAGATTTTG 660
    |||
    |||
    |||
Db 205 AlaValAspGluTrpGlyAlaTrpLeuMetValAspGluAlaHisAlaThrGlyLeuPhe 224
    :
    :
    :
QY 661 GCGGATTCGGGACAGAAACGAGTGAATACTTTGGTGTGT-----CCCGACATTGTT 714
    |||
    |||
    |||
Db 225 GlyAspSerGlyGlyValValGlnArgGluGlyLeuSerHisArgValAspIleGln 244
    :
    :
    :
QY 715 ATCGGACCTTAACAAAGCTGTTGGCGGAGGAGGTTTGGCGGAGATCAGCGGTC 774
    :
    :
    :
Db 245 LeuGlyThrLeuSerLysAlaLeuAlaSerGlnGlyGlyTyrIleAlaGlyAspGluVal 264
    |||
    |||
    |||
QY 775 TTTATCGACTTTTGTGTAACCATGCCAGAACATTTATCTTTCAAAACCGCTATTCCGCCA 834
    |||
    |||
    |||
Db 265 LeuIleGluTrpLeuLeuSerHisAlaArgSerPheValPheSerThrGlyLeuSerPro 284
    :
    :
    :
QY 835 GCGAGCTGTGCGGCTGCTCAGAGGCTTTCAACATCATTTGAAGCCAGCAGGAAACGA 894
    |||
    |||
    |||
Db 285 ProAsnAlaAlaAlaCysGluAlaLeuArgIleAlaArgLysThr---AspArgAla 303
    :
    :
    :
QY 895 CAGCTTTTATTTCTTATATCAGATGATCAGAACCATCTGAAGAAATATGGTTATGTG 954
    |||
    |||
    |||
Db 304 AlaGluLeuTrpAspThrValAlaThrLeuArgAspGlyLeuGluThrMetGlyTyrGlu 323
    |||
    |||
    |||
QY 955 GTGAAGAGGATCACACACGATTTATTCGTAGTCAATTCGCGATGCCATAAAACGGTC 1014
    |||
    |||
    |||
Db 324 ValLeuGlyGlu---ThrHisIleLeuProValIleValGlyAspArgGlyAspAlaLeu 342
    :
    :
    :
QY 1015 CTATTTGCTGAAACATGCGAGGCAAGGAATTTATCTCTGCAATTCGGCGCGCAAC 1074
    |||
    |||
    |||
Db 343 GluLeuAlaAspArgLeuArgAspHisGlyIleValAlaProAlaIleArgProThr 362
    |||
    |||
    |||
QY 1075 GTTGGCGCGGTTGAACCGCGATTCGA 1101
    |||
    |||
    |||
Db 363 ValProGluGlyThrSerArgIleArg 371
    |||
    |||
    |||
RESULT 9
US-60-752-355-36490
; Sequence 36490, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 36490
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Trichodesmium erythraeum IMS101
US-60-752-355-36490
Alignment Scores:
Pred. No.: 2,07e-62 Length: 393
Score: 721.50 Matches: 162
Percent Similarity: 62.4% Conservative: 85
Best Local Similarity: 40.9% Mismatches: 128
Query Match: 13.1% Indels: 21
DB: 8 Gaps: 7
US-10-681-086-1 (1-3156) x US-60-752-355-36490 (1-393)
QY 34 TCCTGGTTAAACGAGCGGTTAGACAGAAATGAAAGACCGCGGTACATCGTAAACCTGCGG 93
```

```

    |||
    |||
    |||
Db 7 SerTrpIleGluSerSerLeuSerThrIleHisLeuAlaAsnTrpTyrArgSerValLys 26
    :
    :
    :
QY 94 TCAATCGATGAGCGCGGTTCCAGAGAGGAATATTGATCGCGAAATAACAAACGCTCGG 153
    |||
    |||
    |||
Db 27 ThrIleGluSerIleProGlyAlaIleIleLysLeuGluGlyLysLysLeuLeuAsnPhe 46
    :
    :
    :
QY 154 TCCTCAAAACAATTTAGGGCTCGCAAGCGATAGACGTTTGTGATCGATGAGAGCAACAA 213
    |||
    |||
    |||
Db 47 AlaSerAsnAspTyrLeuGlyLeuAlaGlyAspGluArgLeuIleAlaIleGln 66
    :
    :
    :
QY 214 GCATTCACGAATTTGGGACAGAACGAGCGGTTTACGTTTACGACGAGCAATTCGGTC 273
    |||
    |||
    |||
Db 67 AlaThrLysGluPheGlySerGlySerThrGlySerArgLeuLeuThrGlyHisArgGlu 86
    :
    :
    :
QY 274 TGGCATCAAAAGCTAGAAAAGAGATTGCCAGCTTTAAACTGACAGAGCGGCGCTCTGCT 333
    |||
    |||
    |||
Db 87 IleHisGlnGlnLeuGluArgGluIleAlaLysLeuLysGlnThrGluSerAlaLeuVal 106
    :
    :
    :
QY 334 TTTTCAGCGGTTACTTGGCCAATGTCGGTCTCTTTTCATCTTCCTTGCAGAAAAGAGAT 393
    |||
    |||
    |||
Db 107 PheSerSerGlyTyrLeuAlaAsnIleGlyValIleSerSerValValSerGlnArgAsp 126
    :
    :
    :
QY 394 GTCATTTTAAAGTACACGAGCTCAATCATGCAAGTATGATCGACGCTCCGACCTTTCTAAG 453
    |||
    |||
    |||
Db 127 LeuIleLeuSerAspGluTyrAsnHisSerSerLeuLysAsnGlyAlaIleLeuSerGly 146
    :
    :
    :
QY 454 GCTGATACAGTTCTTATCGGCATATTGATATCAATGATCTTGAAACCAAGCTGAATGAA 513
    |||
    |||
    |||
Db 147 AlaLysIleIleGluTyrSerHisAsnAsnIleGluTyrLeuLysAsnLysLeuGln 166
    :
    :
    :
QY 514 ---ACACAGCGTTATCAGCGCGCTTTTATCGTAACAGACGAGTATTCACGATGGATGGC 570
    |||
    |||
    |||
Db 167 LysArgGluAsnTyrArgArgSerLeuIleIleThrAspSerValPheSerMetAspGly 186
    :
    :
    :
QY 571 ACAATCGCCCTCTTGCATCAGATCATCTCACTCGGAAACGCTATCATGCTTCGTCGTC 630
    |||
    |||
    |||
Db 187 AspLeuCysLysLeuProLeuLeuLeuAspLeuAlaGluLysTyrAsnSerMetLeuLeu 206
    :
    :
    :
QY 631 GTTGATGATGCCCCACCAACAGAGATTTTGGGGGATTTCGGGCAACAGGAACGAGTGAATAC 690
    |||
    |||
    |||
Db 207 ValAspGluAlaHisAlaThrGlyValPheGlyIleAsnGlyGlyCysValGluHis 226
    :
    :
    :
QY 691 TTTGGTGTGTTGCCCGAC-----ATTGTTATCGCACCTTAAAGCAAGCTGTGTGGC 741
    |||
    |||
    |||
Db 227 PheAsn---CysThrGlyLysGlnLeuIleGlnIleGlyThrLeuSerLysAlaLeuGly 245
    :
    :
    :
QY 742 GCGGAAGGAGGTTTTCGGGCGAGATCAGCGGCTTTCATCGACTTTTGTGTAACCATGCC 801
    |||
    |||
    |||
Db 246 SerLeuGlyGlyTyrValAlaGlySerLysAsnLeuIleGluPheLeuArgAsnArgThr 265
    :
    :
    :
QY 802 AGAACATTTATCTTCAACCGCTATTTCGCCAGCCAGCTGTGCGGCTGCTCACAGGCT 861
    |||
    |||
    |||
Db 266 ProThrTrpIleIleTyrThrGlyLeuThrProAlaAspThrAlaAlaLeuThrAla 285
    :
    :
    :
QY 862 TTCACATCATTTGAAGCCAGACGAGGAAACACGACAGCTTTTATTTCT----- 909
    |||
    |||
    |||
Db 286 IleLysIleIleLysLysGluProGluArgArgMetLysLeuTrpGlnAsnLeuGluIle 305
    :
    :
    :
QY 910 TATATCAGCATGATCAGAAC-----AGTCTCAAGAAATATGGTTATGTGTTGTAAGGA 963
    |||
    |||
    |||
Db 306 PheIleAsnLeuLeuGluThrGluSerGlnLeuLeuHisLysGlyLysThrSerAsn 325
    :
    :
    :
QY 964 GATCACACACCGATTATTCCTGTAGTCATTGGCGGATGCCCATAAACGGCTCTATTGCT 1023
    |||
    |||
    |||
Db 326 TyrGluSerProIleIleCysPheProLeuLysAsnAlaValGluAlaLeuLysValGly 345
    :
    :
    :
QY 1024 GAAAACTGCAGGCAAGGAATTTATGCTCTGCCATTCGGCCGCGCAACCGTTGCGCCG 1083
    |||
    |||
    |||
Db 346 GluLysLeuLysGlnGluGlyIlePheAlaProAlaIleArgProThrValAsnThr 365
    :
    :
    :
QY 1084 GGTGAAAGCGGATTCGAAGCTTGGGCGACGAGCTCGAGATCATCGGGAATGATTTATAAA 1143
    |||
    |||
    |||
```

```
Db 366 SerArgileArgilleSerLeuMetSerThrHisGluThrSer----- 379
Qy 1144 TAAAAAAGCACCTGAAAGGTGCTTTTTTTTGTGATGGTTTGAACCTTG 1191
Db 380 -----HisLeuGlnGln---LeuileAlaAlaLeuileAenLeu 391
RESULT 10
US-60-752-355-20828
; Sequence 20828, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20828
; LENGTH: 393
; TYPE: PRT
; ORGANISM: Thermoplasma volcanium GSS1
US-60-752-355-20828
Alignment Scores:
Pred. No.: 1.8e-61 Length: 393
Score: 712.00 Matches: 153
Percent Similarity: 61.3% Conservative: 70
Best Local Similarity: 42.0% Mismatches: 137
Query Match: 12.9% Indels: 4
DB: 8 Gaps: 3
US-10-681-086-1 (1-3156) x US-60-752-355-20828 (1-393)
Qy 34 TCCTGGTTAAACGAGCGGTAGACAGATGAAGACCGCGGTACATCGTAACCTGCGG 93
Db 5 AsnTrpValAspGluGluileSerAlaLeuLysAlaGluGlyArgTyrValProileArg 24
Qy 94 TCAATGGATGAGCGCGGTCCAGAGGAGTAATGATGGCGGAATCAAAACGGTCTGG 153
Db 25 ThrileGluSerAlaGlnGlyAlaTrpValLysileAenGlyLysGlnValLeuAenMet 44
Qy 154 TCCTCAAAACAATTTATAGGCTCGCAGCGCATAGACGTTTGCATCGAGCCCAACA 213
Db 45 CysSerAenAntyLysLeuGlyPheAlaAenHisProGluThrLysLysAlaAlaileGlu 64
Qy 214 GCATTTGACGAATTTGGACAGGAAGACGCGGTTCACGTTTAAACGACGCAATTCGGTC 273
Db 65 AlaileGluLutyGlyValGlyAlaGlyAlaValArgSerileAlaGlyThrAspGlu 84
Qy 274 TGGCATGAAGAGCTAGAAAGAAGATTGCCAGCTTTAACTGACAGAGCGGCGCTCTG 333
Db 85 IleHisileLysLeuGluGluLysileAlaSerPheLysHisSerGluAlaAlaLeuVal 104
Qy 334 TTTTCGAGCGGTACTTGGCCAATGTCGGTCTCTTTTCATCTCTGCCAGAAAGAGAT 393
Db 105 TyrGlnGlyGlyLeuLeuAlaAenLeuGlyThrValProAlaLeuValGlyLysAspAsp 124
Qy 394 GTCATTTTAAGTACCAGCTCAATCATGCAAGTATGATCGACGGTGGCGGCTTTCTTAAG 453
Db 125 IleilePheSerGluGluLeuAenHisAlaSerileileAspGlyThrArgLeuSerPro 144
Qy 454 GCTGATCAGTTGTTATCGCATATTGATGATGATGATGATGATGATGATGATGATGAT 513
Db 145 AlaLysArgileValTyrLysHisLeuSerValAspAspLeuGluLysGlnAlaLysGlu 164
Qy 514 ACACAGCGTTAT---CACGCGCGTTTATCGTAACACAGCGGAGTATTTCAGCATGGATGGC 570
Db 165 AsnArgSerSerGlyLysLeuAlaLeuValilleThrAspGlyValPheSerMetAspGly 184
Qy 571 ACAATCGCCCTCTTGTATCAGATCATCTCATCTGCGAAACCGCTATCATGCTTCGTGGTC 630
Db 185 AspileAlaProLeuProGluileValAspAlaGluLysTyrAspMetThrTyr 204
```

```
Qy 631 GTTGATGATCCACGCAACAGGAGTTTGGCGGATTCGGGACAAGCAAGCAGTGAATAC 690
Db 205 ValAspAspAlaHisGlyGluGlyValLeuGlyAspHisGlyArgGlyValLeuAsnTyr 224
Qy 691 TTTGGTGTT-----TGTCCCGCAGCATTTGATCGGCACCTTTAAGCAAAAGCTGTTGGCGCG 744
Db 225 PheHisLeuGluAspLysValAspIleGluMetGlyThrPheSerLysAlaLeuGlySer 244
Qy 745 GAAGGAGGTTTGGCGGAGGATCAGCGGTCTTCATCGACTTTTGTGTAACCATGCCAGA 804
Db 245 MetGlyGlyPheValAlaGlySerAlaGluMetIleAspLeuLeuLysGlnLysAlaAarg 264
Qy 805 ACATTTATCTTCAACCGCTATTCCCGCAGCCAGCTGTGGCTGTCTCAGCAGGCTTTC 864
Db 265 ProPheLeuPheSerSerAlaLeuAenProGlyAspAlaAlaValLeuLysAlaIle 284
Qy 865 AACATCATTTGAAGCCAGCAGGAAAAACAGACGCTTTTATTTTATATATATCAGCATGATC 924
Db 285 GluileLeuGluLysAspSerLeuValLysLysLeuTrpGluAenAlaAspLeuLeu 304
Qy 925 AGAACCATGCTGAAGATATGGGTTATGTTGGTGAAGAGATCACACACCATTTATTCCT 984
Db 305 LysLysSerLeuAlaAspLeuGlyTyrAsnThrGlyHisSerLysThrProileThrPro 324
Qy 985 GTAGTCATTTGGCGATGCCCATAAACGGTCTATTGCTGTAATAAACTG---CAGGGCAAG 1041
Db 325 ValilleGlyAspGluLysLysThrValGluLeuSerThrLysLeuTyrGluGluLys 344
Qy 1042 GGAATTTATGCTCTGCTCCATTCGGCCGCCCAACCGTTCGGCGGTGAAGCCGATTCGA 1101
Db 345 AsnValPheAlaSerProileValTyrProThrValProLysGlyThrAlaAargileArg 364
Qy 1102 AGCTTGGCGCAGC 1113
Db 365 LeuMetProSer 368
RESULT 11
US-60-752-355-13501
; Sequence 13501, Application US/60752355
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark et al.
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits
; FILE REFERENCE: 38-21(53720)B
; CURRENT APPLICATION NUMBER: US/60/752,355
; CURRENT FILING DATE: 2005-12-21
; NUMBER OF SEQ ID NOS: 52803
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 13501
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Thermoplasma volcanium GSS1
US-60-752-355-13501
Alignment Scores:
Pred. No.: 1.81e-61 Length: 396
Score: 712.00 Matches: 153
Percent Similarity: 61.3% Conservative: 70
Best Local Similarity: 42.0% Mismatches: 137
Query Match: 12.9% Indels: 4
DB: 8 Gaps: 3
US-10-681-086-1 (1-3156) x US-60-752-355-13501 (1-396)
Qy 34 TCCTGGTTAAACGAGCGGTAGACAGATGAAGACCGCGGTACATCGTAACCTGCGG 93
Db 8 AsnTrpValAspGluGluileSerAlaLeuLysAlaGluGlyArgTyrValProileArg 27
Qy 94 TCAATGGATGAGCGCGGTTCACAGAGGAGTAATGATGGCGGAATCAAAACGGTCTGG 153
Db 28 ThrileGluSerAlaGlnGlyAlaTrpValLysileAenGlyLysGlnValLeuAenMet 47
Qy 154 TCCTCAAAACAATTTATGAGGCTCGCAGCGCATAGACGTTTGCATCGAGCCCAACA 213
```



Db 48 CysSerAsnAsnTyrLeuGlyPheAlaAsnHisProGluThrLysLysAlaAlaIleGlu 67  
QY 214 GCATTCACCAATTTCCGACAGCAAGCAGCGGTTACAGTTTAAACGACGCAATTCGGTC 273  
Db 68 AlaIleGluGluTyrGlyValGlyAlaGlyAlaValArgSerIleAlaGlyThrAspGlu 87  
QY 274 TGCATGAAAGCTAGAAAAGAGATTGCCAGCTTTTAAACTGACAGAACGCCCTCGTC 333  
Db 88 IleHisIleLeuGluGluGluLysIleAlaSerPheLysHisSerGluAlaLeuVal 107  
QY 334 TTTTCAGCGGTACTTGGCCATGTCGGTCTCTTCATCTCCAGCAAGAGAGAT 393  
Db 108 TyrGlnGlyLeuLeuAlaLeuLeuGlyThrValProAlaLeuValGlyLysAspAsp 127  
QY 394 GTCATTTTAAGTACACGAGCTCAATCATGCAAGTATGATGACGCGTCCGACCTTCTAAG 453  
Db 128 IleIlePheSerGluGluLeuAsnHisAlaSerIleAlaSerGlyThrArgLeuSerPro 147  
QY 454 GCTGATACAGTTGTTTATCGGCATATTGATGATGATCTTGAACAAGCTGTAATGAA 513  
Db 148 AlaLysArgIleValTyrLysHisLeuSerValAspAspLeuGluLysGlnAlaLysGlu 167  
QY 514 ACACAGCGTTAT---CAGCGCGCTTTTATCGTAACAGCGAGTATTCAGCATGCGTC 570  
Db 168 AsnArgSerSerGlyLysLysAlaLeuValIleThrAspGlyValPheSerMetAspGly 187  
QY 571 ACATCGCCCTCTGATCAGATCAGTCTCATCTGCGAAGCGCTATCATGCTTCGTCGTC 630  
Db 188 AspIleAlaProLeuProGluIleValAspIleAlaGluLysTyrAspIleMetThrTyr 207  
QY 631 GTTGATGATGCCCAACAGCAGAGTTTGGCGGATTTCGGGCAACAGGAACGAGTGAATAC 690  
Db 208 ValAspAspAlaHisGlyGluGlyValLeuGlyAspHisGlyArgGlyIleValAsnTyr 227  
QY 691 TTTGGTGT---TGTCCCGACATTTATCGGCACCTTAAGCAACAGCTGTTCGCGCG 744  
Db 228 PheHisLeuGluAspLysValAspIleGluMetGlyThrPheSerLysAlaLeuGlySer 247  
QY 745 GAAGGAGTTTTCGGCAGGATCAGCGGTCTTCATCGACTTTTGTGAACCATGCCAGA 804  
Db 248 MetGlyGlyPheValAlaGlySerAlaGluMetIleAspLeuLeuLysGlnLysAlaArg 267  
QY 805 ACATTTATCTTTCAACCGCTATTTCGCCAGCCAGCTGTGCGGCTGTCTCACGAGGCTTC 864  
Db 268 ProPheLeuPheSerSerAlaLeuAsnProGlyAspAlaAlaValLeuLysAlaIle 287  
QY 865 ARAATCATGAAGCCAGCAGGAAAAACGACAGCTTTTATTTTATTCATCAGCATGATC 924  
Db 288 GluIleLeuGlyLysAspSerLeuValLysLysLeuTrpGluAsnAlaAspLeuLeu 307  
QY 925 AGAACGATCTGAAGATATGGGTATGTCGTGNAAGGAGATCACACACCGATTTCTCT 984  
Db 308 LysLysSerLeuAlaAspLeuGlyTyrAsnThrGlyHisSerLysThrProIleThrPro 327  
QY 985 GTAGTCAATTCGGCATGCCATAAAGCTCTATTGCTGAAACAACTG---CAGGGCAAG 1041  
Db 328 ValIleIleGlyAspGluLysLysThrValGluLeuSerThrLysLeuTyrGluLys 347  
QY 1042 GGAATTTATCTCTGCGATTCGGCCCGCAACCGTTGCGCGCGGTTGAACCGGATTCGA 1101  
Db 348 AsnValPheAlaSerProIleValTyrProThrValProLysGlyThrAlaArgIleArg 367  
QY 1102 AGCTTGGCGCAGC 1113  
Db 368 LeuMetProSer 371

## RESULT 12

US-60-752-355-4236  
; Sequence 4236, Application US/60752355  
; GENERAL INFORMATION:  
; APPLICANT: Abad, Mark et al.  
; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits  
; FILE REFERENCE: 38-21(53720)B

; CURRENT APPLICATION NUMBER: US/60/752,355;  
; CURRENT FILING DATE: 2005-12-21  
; NUMBER OF SEQ ID NOS: 52803  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 4236  
; LENGTH: 452  
; TYPE: PRT  
; ORGANISM: Thermus thermophilus HB27  
US-60-752-355-4236

Alignment Scores:  
Pred. No.: 1 2e-60 Length: 452  
Score: 704.00 Matches: 155  
Percent Similarity: 60.5% Conservative: 72  
Best Local Similarity: 41.3% Mismatches: 142  
Query Match: 12.8% Indels: 6  
DB: 8 Gaps: 4

US-10-681-086-1 (1-3156) x US-60-752-355-4236 (1-452)

QY 1 GGATCCACGAGGTAA---CGAGCCTTTGAAGATTGAT-----TCCTGGTTAAACGAGCGG 51  
Db 50 GlyProArgArgLeuAspLysAlaMetSerLeuAspLeuArgAlaArgValArgAspGlu 69  
QY 52 TTAGACAGAAATGAAGAAGCGCGGTACATCGTAACCTCGGTCATGATGATGAGAGCGCG 111  
Db 70 LeuGluArgLeuLysArgGluGlyLeuTyrIleSerProLysValLeuGluAlaProGln 89  
QY 112 GTTCCAGAGAGNATATTGATCGCGAANAATCAACCGTCTGCTCCCAACAATTTTAA 171  
Db 90 GluProAlaThrArgValGluGlyArgGluValValAsnLeuAlaSerAsnTyrLeu 109  
QY 172 CGGCTCGCAAGCGATAGACGTTTGTGATCGATCGAGCCCAACAGCATTCGACCAATTTGG 231  
Db 110 GlyPheAlaAsnHisProTyrLeuLysGluLysAlaArgGlnTyrLeuGluLysTrpGly 129  
QY 232 ACAGGAAGCAGCGGTTTCAGCTTTAAACGACGAGCAATTCGGTCTGGCATGAAAGCTAGAA 291  
Db 130 AlaGlySerGlyAlaValArgThrIleAlaGlyThrPheThrTyrHisValGluLeuGlu 149  
QY 292 AAGAAGATTGCCAGCTTTAAACTGACAGAGCGCGCTGCTGTTTCGAGCGTTACTTG 351  
Db 150 GluAlaLeuAlaArgPheLysGlyThrGluSerAlaLeuValLeuGlnSerGlyPheThr 169  
QY 352 GCCAATGTCGCTGCTCTTCATCTTCCTCCAGAAAAGAGAGATGCTCATTTTAAGTGACCAG 411  
Db 170 AlaAsnGlnGlyValLeuGlyAlaLeuLeuLysGluLysValValPheSerAspGlu 189  
QY 412 CTCATTCATCAAGTATGATCGACGCTGCCGACTTTCTAAGGCTGATACAGTTGTTTAT 471  
Db 190 LeuAsnHisAlaSerIleIleAspGlyLeuArgLeuThrLysAlaThrArgLeuValPhe 209  
QY 472 CGGCATATTGATATGATGATCTTGAACACAGCTGAATCAACACAGCGTTATCAGCGC 531  
Db 210 ArgHisAlaAspValAlaHisLeuGluLeuLysAlaHisAspThrAspGlyLeu 229  
QY 532 CGTGTATTTCGTAAACAGCGAGTATTTCAGCATCGATGGCAACAATCGCCCTCTTGATCAG 591  
Db 230 LysLeuIleValThrAspGlyValPheSerMetAspGlyAspIleAlaProLeuAspLys 249  
QY 592 ATCATCTCACTTCGAAACGCTATCATGCTTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 651  
Db 250 IleValProLeuAlaLysLysTyrArgAlaValValTyrValAspAspAlaHisGlySer 269  
QY 652 GGAGTTTTCGGCGATTTCGGGCAAGCAAGAGTGAATCTTTCGTTGTT-----TGTCCC 705  
Db 270 GlyValLeuGlyGluLysGlyLysGlyThrValHisPheGlyPheHisGlnAspPro 289  
QY 706 GACATTGTT---ATCGGCACCTTTAAGCAAGCTGTTGGCGCGGAAAGAGGTTTTTCGGCA 762  
Db 290 AspValIleGlnValAlaThrLeuSerLysAlaTrpAlaGlyIleGlyGlyTyrAlaAla 309  
QY 763 GGATCAGCGGCTCTTCATCGACTTTTTCGTAACCATGCGCAGAACATTTATCTTCAAACC 822



Db 322 ThrProIleThrProValLeuPheGlyGluAlaProLeuAlaPheGluAlaSerArgLeu 341  
 QY 1030 CTGACGGCAGGAATTTATGTCTCTGCCATTGGCCGCAACCGTTGCGCCGGTGAA 108  
 Db :|||:||||:||||| ||| ||||| |||:|  
 Db 342 LeuLeuGluGluGlyValPheAlaValGlyIleGlyPheProThrValProArgGlyLys 361  
 QY 1090 AGCGGATTCGAAGCTTG 1107  
 Db :|||:|||||:||||| :|||:|  
 Db 362 AlaArgIleArgAenIle 367  
 RESULT 15  
 US-60-752-355-35003  
 ; Sequence 35003, Application US/60752355  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Abad, Mark et al.  
 ; TITLE OF INVENTION: Transgenic plants with enhanced agronomic traits  
 ; FILE REFERENCE: 38-21(53720)B  
 ; CURRENT APPLICATION NUMBER: US/60752,355  
 ; CURRENT FILING DATE: 2005-12-21  
 ; NUMBER OF SEQ ID NOS: 52803  
 ; SOFTWARE: PatentIn version 3.3  
 ; SEQ ID NO 35003  
 ; LENGTH: 393  
 ; TYPE: PRT  
 ; ORGANISM: Thermoplasma acidophilum DSM 1728  
 US-60-752-355-35003  
 Alignment Scores:  
 Pred. No.: 2,78e-60 Length: 393  
 Score: 700.00 Matches: 154  
 Percent Similarity: 59.9% Conservative: 64  
 Best Local Similarity: 42.3% Mismatches: 142  
 Query Match: 12.7% Indels: 4  
 DB: 8 Gaps: 3  
 US-10-681-086-1 (1-3156) x US-60-752-355-35003 (1-393)  
 QY 34 TCTGTGTTAAACGACGGTTAGACAGAAATGAAAGACCGCGGTACATCGTAACTCGGG 93  
 Db |||||:|:| ||||| ||||| :|||:|  
 Db 5 SerTrpValGluGluLeuLeuSerAlaLeuLysAlaGluGlyArgTyrValProIleArg 24  
 QY 94 TCAATGATGGAGCGCGGTTCCAGACAGAGATATTGATGCGGAAAAATCAACGGCTCGG 153  
 Db :|||:|:| ||||| ||||| :|||:|  
 Db 25 ThrIleGluSerAlaGlnGlySerTrpValThrIleGlyLysLysValLeuAsnMet 44  
 QY 154 TCCTCAAAACAATTTATTTAGGCGTCGCAAGCGATAGAGCTTTTGATCGATGCAGCCCAACA 213  
 Db |||||:|:| ||||| ||||| :|||:|  
 Db 45 CysSerAsnAsnTyrLeuGlyPheAlaAsnHisProGluThrLysLysAlaAlaIleGlu 64  
 QY 214 GCATTGACGCAATTTGGGACAGGAAGCAGCGGTTTCAGCTTTTAAACGACAGCAATTCGGTC 273  
 Db |||||:|:| ||||| ||||| :|||:|  
 Db 65 AlaIleGluGlnTyrGlyValGlyAlaGlyAlaValArgSerIleAlaGlyThrAspGlu 84  
 QY 274 TGGCATGAAAGCTAGAAAAGAGATTGCCAGCTTTAACTGACAGAACGCCGCTCTGCTG 333  
 Db |||||:|:| ||||| ||||| :|||:|  
 Db 85 IleHleAlaArgLeuGluGluLysIleAlaLysPheLysHisMetGluSerAlaLeuVal 104  
 QY 334 TTTTCGAGCGGTACTTGGCCAATGTCGGGTGCTCTTCATCCTTCCAGAAAAGGAACAT 393  
 Db :|||:|:| ||||| ||||| ||||| :|||:|  
 Db 105 TyrGlnGlyLeuLeuAlaAsnValGlyThrIleProAlaLeuValGlyLysAspAsp 124  
 QY 394 GTCAATTTTAAGTGACCAAGCTCAATCATCGCAAGTATGATCGACGCTGCCGACTTTCTAAG 453  
 Db |||||:|:| ||||| ||||| ||||| :|||:|  
 Db 125 ValIlePheSerGluGluLeuAsnHisAlaSerIleIleAspGlyThrArgLeuSerSer 144  
 QY 454 GCTGATACAGTTGTTATCGGCATATTGATATGATGATCTTGAACCAACAGCTGAATGAA 513  
 Db |||||:|:| ||||| ||||| ||||| :|||:|  
 Db 145 AlaLysArgIleValTyrLysHisLeuSerValGluAspGluLysGlnIleArgGlu 164  
 QY 514 ACACAG---CGTTATCAGCCCGTTTTTATCGTAAACAGACGGAGTATTCAGCATGATGCG 570  
 Db :|||:|:| ||||| ||||| ||||| :|||:|  
 Db 165 AsnArgSerAlaPheLysLysAlaLeuValIleThrAspGlyValPheSerMetAspGly 184

```
Qy 571 ACAATCGCCCTCTTGATCAGATCATCTCACTTGGAAACGCTATCATGCTCTGCTGTC 630
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
185 AspileAlaProLeuProGluIleThrGluValAlaGluLysAsnAspValMetValTyr 204
Qy 631 GTTGATGATGCCACGCAACAGGAGTTTGGCGGATTCGGGCAACAGGAACGAGTGAATAC 690
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
205 ValAspAspAlaHisGlyGluGlyValLeuGlyAspHisGlyArgGlyIleValAsnTyr 224
Qy 691 TTTGGTGT-----TGTCGCCGACATGTTATCGGCACCTTAAGCAAGCTTGTGGCGC 744
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
225 PheHisLeuGluAspArgValAspileGluMetGlyThrPheSerLysAlaLeuGlySer 244
Qy 745 GAAGGAGGTTTGGCGGAGGATCAGCGCTCTTCATCGACTTTTGTGCTGAACCATGCCAGA 804
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
245 MetGlyGlyPheValAlaGlySerAlaAspLeuIleAspLeuLeuLysGlnLysAlaArg 264
Qy 805 ACATTTATCTTTCAAAACCGCTATTCCGCCACCGCTGTGCGGCTGCTCACGAGCTTTC 864
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
265 PropheLeuPheSerSerAlaLeuAsnProGlyAspAlaAlaValLeuLysAlaIle 284
Qy 865 AACATCATTTGAAGCCAGCAGGAAAAACGACAGCTTTTATTTCTTATATCAGCATGATC 924
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
285 GluIleLeuGluLysAspSerLeuIleLysLysLeuTrpHisAsnSerAspileLeu 304
Qy 925 AGAACCACTCTGAAGAATATGGGTTATGTGTGTAAGGAGATCACACACGATTATTCTCT 984
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
305 LysLysSerLeuSerAspValGlyTyrAsnThrGlyHisSerLysThrProIleThrPro 324
Qy 985 GTAGTCATTTGGCGATGCCCATAAACCGTCTTATTTGCTGAAAAACTG---CAGGGCAAG 1041
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
325 ValIleIleGlyAspGluLysLysThrValGluLeuSerLysMetLeuTyrAspGluLys 344
Qy 1042 GGAATTTATGCTCTGCCATTTCGCCGCCCAACCGTTGGCCGGGTGAAGCCGGATTGCA 1101
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
345 AspValPheAlaSerProIleValTyrProThrValProLysGlyThrAlaArgIleArg 364
Qy 1102 AGCTTGGGCAGC 1113
Db |||||
365 LeuMetProSer 368
```

Search completed: February 10, 2006, 03:51:30

Job time : 93 secs